"High Speed All Optical Networks"

Annual Report

Work performed by

Imrich Chlamtac, (413) 545-0712
Aura Ganz, (413) 545-0574
Department of Electrical and Computer Engineering
University of Massachusetts
Amherst, MA 01003

During the period May 1, 1989 through April 30, 1990

For

NASA AMES Research Center
Mofsett Field, CA 94035

Table of Contents

WIDE AREA NETWORK

Report Summary

An inherent problem of conventional point-to-point WAN architectures is that they cannot translate optical transmission bandwidth into comparable user available throughput due to the limiting electronic processing speed of the switching nodes. This report presents the first solution to WDM based WAN networks that overcomes this limitation. The proposed Lightnet architecture takes into account the idiosyncrasies of WDM switching/transmission leading to an efficient and pragmatic solution. The Lightnet architecture trades the ample WDM bandwidth for a reduction in the number of processing stages and a simplification of each switching stage, leading to drastically increased effective network throughputs.

The principle of the Lightnet architecture is the construction and use of virtual topology networks, *embedded* in the original network in the *wavelength* domain. For this construction Lightnets utilize the new concept of lightpaths which constitute the links of the virtual topology. Lightpaths are all-optical, multihop, paths in the network that allow data to be switched through intermediate nodes using high throughput passive optical switches. The use of the virtual topologies and the associated switching design introduce a number of new ideas: WHICH ARE DISCUSSED IN DETAIL

1. The use of embedded *regular* topologies reduces the average number of active processing stages a packet has to traverse in the network. With a smaller number of stages, the number of service instances per packet is reduced, so that the total number of packets that can be processed in the network per unit of time, i.e. the network throughput, is increased. Certain regular topologies, furthermore provide inherent load balancing, leading to reduced buffering requirements. Most regular topologies also entail simplification of network control procedures, such as routing, thus further reducing the complexity of network protocols.

2. The construction of the regular topologies in a *virtual mode* provides a feasible approach for establishing and maintaining regular structures in wide area networks, which due to distance and cabling considerations are characterized by arbitrary topologies.

3. Lightnets introduce a two level switching and distribute the processing/switching requirements between the electronic and optical switching capabilities of the intermediate nodes according to the capability of each. Transmissions within lightpaths use passive optical WDM switches whose switching bandwidth matches the rates of optical links. Transmissions between lightpaths use active electronic switches residing in the nodes of the virtual topology. Thus only a *fraction* of total data needs to be switched *actively* at each intermediate node, so that the effective link throughput is no longer limited by the "electronic switching bandwidth". In this way, Lightnets can overcome the electronic switching/processing bandwidth bottleneck of intermediate nodes leading to an effective network throughput that can be higher by an order of magnitude than in current wide area networks.

Quantitative results derived so far, support the performance expectations of the proposed Lightnet architecture.

1. Introduction

Current network architectures fail to meet the integrated demands of emerging communication applications. First and foremost, a substantial increase in network bandwidth must be provided to support applications such as HDTV, super-computer communications and video-conferencing [1,4]. Co-existing with these vast bandwidth consumers, there will continue to be applications with substantially smaller requirements. Thus, in addition to the need for high bandwidth, a bandwidth dynamic range of up to seven orders of magnitude must be contended with *efficiently*. Reliability and availability will also become critical issues in future high speed networks carrying services previously supported by different networks. Clearly, the degree of reliability of the new network must be at least as high as that provided in the past by the network carrying the most stringent of the integrated applications.

Wavelength division multiplexing is a leading technique for providing the very high bandwidth needed in the emerging communication environments [2-3]. Conventional WAN architectures cannot, however, utilize the entire bandwidth offered by WDM due to the growing discrepancy between the electronic processing rates and the optical transmission rates [4]. In fact, packet switching solutions have traditionally been motivated by the need for efficient utilization of bandwidth at the expense of increased processing in the nodes. The performance potential of these networks is thus limited by the bottlenecks created by the switching, buffering and processing requirements at these nodes. Today, leading approaches for wide bandwidth WANs continue to be based on packet switching, often termed "fast packet switching" [5,6]. In these solutions, packets are not required to wait and be error checked at intermediate nodes. However, buffering, E/O conversion of the packet header and routing oriented processing are still required. Therefore, with these solutions, the node bottlenecks created by the discrepancy between optical transmission and electronic processing/buffering capabilities are not removed. This leads to networks

with a limited effective throughput, that can sometimes be only a small fraction of the optical transmission bandwidth [1].

The Lightnet architecture presented here addresses this fundamental problem of high speed wide area networks by trading the ample transmission bandwidth provided by WDM, for a reduction in the amount of switching/processing required per end-to-end packet transmission. This reduction is achieved by introducing the concept of a) *virtual regular topologies*, b) the use of all-optical multihop routes - the *lightpaths*, for the construction of these virtual topologies and by c) the development of these concepts in a way which is consistent with WDM transmission/switching trends. Specifically, Lightnets guarantee a reduced number of processing stages per transmission of each packet, the simplification of each switching stage and the shifting of the switching load from active electro/optic switches to passive, preset, photonic switches.

In foreseeable future, passive switches are the only solution capable of delivering switching bandwidth on the order of the WDM transmission rates. On the other hand we observe, that the use of preset, passive switches, leads to increased bandwidth consumption: With present switches, the routes of lightpaths are fixed, so that packets potentially travel longer paths than required by per packet routing in one of the existing packet/circuit switching modes. This leads to the principle of a tradeoff in Lightnet, whereby bandwidth is "sacrificed" for simplifying switching in a way that allows high bandwidth passive optical switches to be utilized. In addition to allowing the use of high throughput passive optical switches, lightpaths permit the construction of virtual topologies, embedded in the original physical topology. In this construction, lightpaths become the links of the virtual topologies. Due to the fact that the regular topologies are virtual, it becomes feasible to *set up* and *maintain* regular topologies in the wide area domain. Furthermore, by selecting certain *regular* topologies as the Lightnet virtual topology the following performance related benefits can be pursued : 1) Inherently load balanced topologies can be built reducing the problem of congestion, thus lowering buffering requirements. 2) Simplified routing and congestion control procedures can be implemented in the active switches further reducing the amount of processing required per packet transmission.

The Lightnet architecture design is associated with the following hardware con-

siderations: 1) The active switch size, determined by the virtual topology node degree and being equal to the number of lightpaths terminating at the node. 2) The passive switch size, determined by the physical node degree, and the number of wavelengths. 3) The number of wavelengths required, associated with the number of transmitters and receivers required, the passive switch size and the transmission technology. By accounting for these hardware considerations, the Lightnet architecture carries not only the potential to provide truly high speed WANs, but also leads to solutions that are practical.

2. Fundamental Considerations in Lightnet Solution

The proposed Lightnet architecture is based on the following observations:

Given:

1) the wavelength division multiplexing is the emerging transmission technique,

2) in photonic switching:

   a) the passive switch bandwidth is comparable to the transmission bandwidth,

   b) the active switch bandwidth lags far behind.

Therefore:

a) To obtain a high throughput WDM based WAN passive (preset) optical switches must be utilized. Preset switches lead to the "lightpath" concept, pre-established all-optical transmission paths between source-destination nodes (not necessary adjacent) in the network.

A lightpath is implemented by using the same wavelength on its route and passive photonic switches. Since no active switching is incurred over a lightpath, its bandwidth is equivalent to the fiber bandwidth.

However:

the number of wavelengths and, from independent considerations, the switching complexity that would be required to establish a lightpath between all source-destination pairs in a WAN, are prohibitively high!

Therefore:

lightpaths will be established between a *subset* of node pairs only.

In Lightnet, we therefore introduce the second basic idea - the use of lightpaths as virtual links for generating new, virtual, topologies. We design the virtual topologies to optimally balance passive and active switching, (the key performance bottleneck of high speed networks) and to provide performance enhancement in a way independent of the application traffic requirements. Specifically, as pointed out in section 1, by selecting *regular* topologies:

The number of active switches a packet will traverse can be reduced logarithmically, compared to the number of active switching stages in the original topology. The virtual topologies can further provide inherent load balancing, simplified routing and congestion control procedures. Finally, it is important to observe that the Lightnet approach constructs the regular topologies virtually so that the benefits of transmission over regular structures can be introduced to any physical topology network. In this respect, the Lightnet architecture is unique not only in the wide area, but also in the local communication domain.

3. Overview of Lightnet Architecture Design Issues

The construction of a new architecture entails a large variety of problems and in that sense Lightnet is no exception. In this report we present, investigate and propose an initial solution to four basic Lightnet design issues:

1) *WHICH* regular topology to embed?

2) *HOW* to embed a regular topology in the physical WAN?

3) How to establish the *LINKS* of the virtual topology, i.e. the lightpaths?

4) How to realize the *NODES* of the virtual network, i.e. the switches?

The answers to these questions have to take into consideration the technological limitations of WDM transmission and switching on one hand, while providing a truly "high speed" performance, on the other.

Choice of Virtual Topology: The topology determines the number of nodes actively involved in transmitting a packet. It thus determines the total processing/buffering required to transmit a packet end-to-end and consequently the effective network throughput. It also determines global issues, such as routing, congestion control, or load balancing.

The choice of topology must also obey the following hardware considerations: 1) The active switch size, determined by the virtual topology node degree and being equal to the number of lightpaths terminating at the node. 2) The passive switch size, determining the number of wavelengths and their assignments to lightpaths. 3) The number of wavelengths required, determining the number of transmitters and receivers required, the passive switch size and the transmission technology [3,4].

Virtual Topology Construction: In Lightnet, a virtual topology is embedded in a general topology by associating the nodes of the virtual topology with the nodes of the original, general topology. The edges of the virtual topology are implemented by creating lightpaths, i.e., utilizing the WDM domain. It can be argued that the best network performance will be obtained by choosing a clique as the Lightnet virtual topology. However, the number of wavelengths that have to be used for this construction and the associated switching complexity are prohibitively high. Therefore, topologies must be constructed through an embedding process which minimizes hardware requirements, e.g., the total number of wavelengths required. Certain topologies, such as regular topologies posses well known performance/control advantages. A polynomial time algorithmic procedure embedding a regular topology graph in a general topology network, while minimizing the number of wavelengths (and consequently the complexity of the associated hardware), is however not known [7,8]. New heuristics must therefore be introduced.

Lightpath Establishment: The *correct* and *efficient* establishment of lightpaths has to be resolved under the wavelengths availability, continuity and switching constraints mentioned above. The correctness aspect of lightpath establishment must avoid the prob-

lems of allocation conflicts mentioned above. For efficiency, for a given set of lightpaths determined by the choice of the virtual topology, the utilization of wavelengths should be maximized. In a dynamic environment, for a given rate of lightpath establishment requests it is necessary to minimize the lightpath blocking probability. Lastly, to allow for feasible, cost-effective switching implementation, a lightpath should be established using a single, identical, wavelength throughout its path. Figures 1a and 1b exemplify the lightpath allocation problem for even a trivial configuration. The figures depict lightpaths establishment in a network with two available wavelengths ($\omega = 2$). In Figure 1a the allocation is done in such a way that any additional future lightpath establishment request can still be accommodated. In the allocation depicted by Figure 1b, if a lightpath request $v_1 \rightarrow v_3$ arrives before an existing lightpath is terminated, it will be blocked.

Lightnet Switching Node Design: In Lightnet a switch has to handle the lightpaths that pass through it (passively) and lightpaths terminating in it (actively). Given the bandwidth limitation of active switches, referred to in the preceding section, the Lightnet switching configuration is targeted towards handling most of the data through the passive switch. The resulting Lightnet switch is composed of three main switching components, the passive switch, the lightpath terminating switch (LTS), and the active switch, as shown in figure 2. The passive switch provides the "intra-lightpath" switching for the lightpaths passing through the node and switches the lightpaths terminating at the node to the active switch via the LTS. The active switch connects the local node to the network and performs "inter-lightpath" switching according to the destinations of the packets arriving on the lightpaths initiating/terminating at the local node. Lightpaths initiating at the local node proceed from the active switch, through the LTS, to the passive switch. In case of large geographical distances, or a large number of intermediate passive switches, all signals departing the passive switch can be regenerated (as shown in figure 2) and subsequently multiplexed according to the output links. By regenerating the signals, the restriction on the lightpath *span*, i.e. the number of hops a lightpath can traverse, can be eliminated.

The following sections deal in detail with each one of the four issues presented above.

4. Virtual Topologies and Performance Potential

In this section, we present the performance potential obtained by a number of regular Lightnet topologies. The performance measures chosen for this comparison in this section are those that are *independent* of the embedding process and of the underlying physical topology. These are the network capacity, the number of buffers per node and the average delay spent by the packets at the intermediate nodes. We note that the the embedding process and the underlying physical topology will, on the other hand, determine the delay spent by the packets on the links, i.e. the propagation delay.

In the first subsection, we describe the regular topologies studied. A performance comparison between a Lightnet regular topology and the underlying physical WAN topologies using store and forward communication is presented in subsections 4.2 and 4.3 respectively.

4.1 Description of Potential Lightnet Topologies

In this subsection, we present the following regular topologies: binary tree, hyper-tree, diamond, hyper-cube, De-Bruijn and torus. For each topology in addition to a brief description, its diameter and node degree are specified. Recall that the diameter of the virtual topology determines the average number of active switches (virtual topology nodes) a packet will traverse in the Lightnet. The virtual node degree represents the number of lightpaths that will terminate at the node, i.e., the degree of the active switch.

Binary Tree: The diameter (maximal number of hops between any two nodes in the network) obtained by a tree is given by $D = 2(\log_2 n - 1)$.

The basic drawback of a tree topology is the congestion arising in the root. For an n node binary tree, assuming uniformly distributed traffic, approximately $\frac{1}{2}n\lambda$ packets per second will pass through the node, λ being the arrival rate per node. Thus, network capacity C is given by $C = n\lambda_{\max} = 2\mu$ where μ is the service rate per node. Therefore, the binary tree can be considered a viable approach only for lightly loaded networks.

Hypertree: The Hypertree was presented in [9] as an "improved" tree displaying

greater resilience to faults and alleviating the congestion at the root. The idea underlying the Hypertree is connecting nodes to other nodes on the same level in the tree, thus allowing movement of packets across the tree without passing through the root. A sample 15 node Hypertree is presented in figure 3.

In the hypertree, each node is connected to one other node on the same level, resulting in a *constant* degree of 4 (two descendants, an ancestor and a "peer" connection). Nodes are assigned numbers following the "natural" enumeration for binary trees (the immediate descendants of node i are $2i$ and $2i+1$, 1 being the root). Peer connections are made between nodes whose corresponding bit representations differ by a single bit. Among the possible connections, the one minimizing network diameter is chosen. A suboptimal bit oriented routing algorithm employing the peer connections in the Hypertree is presented in [9]. The diameter of the Hypertree is $O(\log n)$ but with a lower constant (half the diameter of a simple binary tree).

Diamond: A Diamond consists of two inverted trees connected through the leaves, coupled with peer connections forming a ring at each layer of the tree. Connections also exist between the roots and the first descendants of the root (see figure 4 for a 14 node Diamond). The diameter of an Diamond with $n-2$ nodes is given by $D = 2\log_2 n - 5$, an improvement over the binary tree. However, due to the peer connections as well as direct connections between the roots and the first descendants of the roots, the traffic distribution in an Diamond is much more balanced than that of a binary tree or a Hypertree. The maximal degree of a node in the Diamond is also *constant*, given by 5 for approximately half of the nodes and a degree of 4 for the others.

Torus: A Torus, a grid with connections wrapping around at the edges intuitively seems to be the optimal constant node degree topology in terms of load balancing due to its complete symmetry. Figure 5 depicts a 4×4 Torus. A Torus has diameter $D = \sqrt{n}$ with each node having degree 4 [10].

De-Brujin Network: A De-Brujin network is established following a simple construction rule based on the binary representation of a node's number [11]. Let k be a node number, $0 \leq k \leq 2^d - 1$. The neighbors of k are the nodes whose numbers correspond to:

- $k_1 = k >> 1$
- $k_2 = compl\_most(k_1)$
- $k_3 = k << 1$
- $k_4 = compl\_least(k_3)$

where $<<, >>$ denote left and right shift and *compl_most, comp_least* denote complement operations on the most and least significant bit. Thus, for example, the neighbors of *00010* are *(00001,00000,00100,00101)*. Figure 6 depicts an 8 node De-Brujin network. The De-Brujin network possesses a diameter of exactly $D = \log_2 n$ and a maximal degree of 4.

Hypercube: The hypercube is a multi-dimensional cube. Two nodes are said to be at a distance of one if their corresponding binary representations differ by a single bit. Thus, a hypercube is a structure of $n = 2^k$ nodes for some integer k, where every two nodes at a distance of one are connected by a link. Figure 7 depicts a hypercube of degree 3 containing 8 nodes. The diameter of a hypercube is given by $D = log_2 n$. The node degree, however, is given by $\log_2 n$, making this the only structure in the group whose node degree is not constant.

4.2 Performance Comparison of Potential Topologies

To evaluate the performance potential of different regular topologies to be embedded in Lightnet a simulation was developed. Packet arrivals where assumed to follow a Poisson distribution with uniform source/destination distribution. Packet processing was assumed to take place at the source and destination nodes as well as at every intermediate node in which a switchover between lightpaths occurs. A packet arriving at a node enters instantly a processing server if one is available or joins a common queue if all are busy. Considering the large discrepancy between transmission and processing speeds the former was neglected. The queue service was modeled by 3 servers, each requiring 10 time units to process a packet. Upon terminating service the packet proceeds instantaneously to the next node (assuming high transmission speeds relative to per packet processing/switching

time) where the process is repeated. Simulation results were obtained with 97% confidence levels.

Table 1 presents the performance measures for the regular topologies discussed in the previous section for networks with 128 nodes. Column 1 depicts the network capacity defined as the maximum throughput (load) for which all node queues still exhibit ergodic behavior. The results shown are all normalized relative to the hypercube capacity. Networks with complete symmetry tend to posses high capacities since all nodes saturate simultaneously. However, we note that the De-Brujin network which is non-symmetrical displays higher capacity than the (symmetric) Torus. This is due both to the substantially smaller average path length in the De-Brujin topology leading to less overall processing per packet delivered and the excellent load balancing achieved in this network despite its asymmetry. The hypercube, being both symmetrical and with low average path length displays substantially higher capacity (more than 50% improvement) than both the Torus and the De-Brujin. Among the tree topologies, the Diamond has the highest capacity, nearly 300% higher than the hypertree.

The last three columns depict maximal buffer usage per node for the three aforementioned loads. As expected, buffer usage follows the respective network capacities rising rapidly as the limit of the ergodicity range is approached. Results displayed are normalized to the ones obtained for the hypercube under low load. While the hypercube is clearly the superior topology, the De-Brujin and the Torus tend to demonstrate similar behavior. Again, the Diamond is the superior one amongst the tree topologies.

Thus, we conclude that from a pure performance measure point of view, the hypercube demonstrates clear advantages. The main drawbacks of the hypercube are its large node degree, $log_2 n$ compared with a constant degree in all other topologies, and network growth only in powers of 2 (true also for the De-Brujin network). If node degree is to be maintained *constant*, the De-Brujin network exhibits slightly superior capacity and substantially lower jitter when compared with the second best constant degree topology, the grid.

4.3 Comparison of Lightnet Topologies with Existing Networks

To further establish the potential performance gain associated with embedding the regular topologies using lightpaths, we compared these topologies with a 61 node, 1978 Arpanet topology, as well as with several randomly generated topologies of the same size.

The random topologies were created so as to mimic closely the Arpanet 1978 topology in terms of diameter, mean node degree and actual node degree distribution and overall number of links. The algorithm was then used to create random networks with 16,32,64, and 128 nodes. 10 networks were created for each network size. Results are shown both for the "best" network among the 10 as well as median results. All results are normalized with respect to the hypercube.

To indicate maximum operational loads (i.e. determining overall capacities) we depict in Figure 8 average delay spent by the packets at intermediate nodes, denoted by D, versus the system throughput (S) for networks with c. 64 nodes.

Network capacities as a function of the network size (normalized to hypercube with 16 nodes) are presented in table 2. Four result sets are presented for the random graphs : two representing maximal (*max_rnd*) and median (*med_rnd*) capacity assuming service times of $\mu = 10$ (equal to the regular networks case) and two representing the same values for $\mu = 40$ (*max_rnd4* and *med_rnd4* respectively). As can be seen, the potential improvement of embedding regular topologies in Lightnets can be more than tenfold.

5. Lightpath Establishment

Since lightpaths are the basic building block of Lightnets, their correct and efficient establishment is crucial to the successful implementation of these architectures. *Lightpath* is defined as an "optical communication path" between two (not necessarily adjacent) nodes, established by allocating the *same* wavelength throughout the route of the transmitted data. As a result, transmissions between lightpath endpoints require no processing or buffering at intermediate nodes. On the other hand, the wavelength continuity constraint increases, intuitively, the number of wavelengths needed to establish a given lightpath

set. Since lightpaths constitute the links of the Lightnet topologies, the total number of lightpaths that can be established, as well as any other technology related limitation on their establishment, such as the maximum length (span) of a lightpath, will affect the virtual topology design. To employ WDM to implement a network architecture based on lightpaths, a number of hardware related issues must, therefore, be examined. Specifically, the multiplexing technique and the resulting number of channels made available, the photonic switches needed and the lightpath span.

With wavelength division multiplexing:

- The number of channels available on each link is limited. Current experimental systems are able to carry up to 20 channels, each modulated at 2 Gbit/s [22]. Similar devices approaching 60 channels are considered feasible in the near future [22,23].

- The design of the photonic switch required for switching lightpaths at the intermediate nodes is closely related to the wavelength continuity property of the lightpaths. Since a lightpath maintains the same wavelength throughout its span, a channel incoming on one wavelength need not be switched to another wavelength. Consequently, in realizing the photonic switch, it is possible to *group* the channels according to their wavelength prior to switching. The photonic switch can thus consist of ω switching matrices, one for each wavelength.

- With regard to the end-to-end lightpath span, experiments conducted recently have successfully demonstrated applications of optical amplifiers [15-19]. For instance, in [15], 25 optical amplifiers were used in series (an amplifier every 80km) to provide transmission of a 2.5 Gbit/s optical signal over 2,223 km of single mode fiber, with a power penalty due to accumulated noise of only 4.2dB.

- Finally, lightpath implementation at intermediate nodes requires the availability of suitable photonic switches. So far, emphasis on switching for LAN and WAN operation concentrated on switch operation with setup rates on par with packet transmission rates, a critical issue for packet switching networks. Electro-optic

switches can be set up in less than 1ns [26], however due to the processing time needed to make the switching decision and, due to their poor crosstalk and attenuation characteristics, they may hinder the performance of an all-optical longhaul implementation. Mechanical optical switches of dimension up to 40x40, can, on the other hand, switch single-mode signals with crosstalk of -90dB and attenuation of around 2dB, thus obtaining crosstalk and attenuation characteristics far better than those offered by electro-optic devices [21]. The use of mechanical switches, despite these characteristics, was not previously considered for data switching networks due to their slow set up speeds. The principle of using preset lightpaths changes this situation dramatically: while the setup time of the mechanical switches is relatively large (50ms) [27], this does not constitute a problem in Lightnet, as lightpaths are not established on a per packet basis and can have lifetimes measured in hours or days.

The combination of these hardware aspects of lightwave communication and the special properties of lightpaths suggest that a lightpath based Lightnet network can present a technologically feasible solution for a wide area wavelength routing network. At the same time practical limits on the size of the photonic switch and on the number of WDM channels per link make the minimization of the number of required wavelengths an important aspect of the Lightnet architecture. We turn to address this problem next.

5.1 Problem Definition and Analysis

Several lightpath establishment policies are motivated by the Lightnet architectures:

A. The first case is the Static Lightpath Establishment policy: a network in which a set of n lightpath requests is predetermined (dictated by the target virtual topology) the objective function is to establish all demands using a minimum number of wavelengths.

B. Since the number of available wavelengths in WDM systems is expected to remain limited, it is of importance to study the problem of establishing a given set of lightpaths when the number of wavelengths, ω, is bounded. We note that in this case,

it is possible that new lightpaths will be blocked. The objective function therefore changes in this case to minimizing the ratio of lightpaths rejected to lightpaths requested, defining the lightpath blocking probability.

C. The interest in studying the case in which lightpaths are established and terminated dynamically, stems from the fact that the Lightnet topology can be modified by reassigning lightpaths. By establishing the lightpaths dynamically, the Lightnet can be reconfigured for purposes of reliability or availability.

D. Last, we consider the problem of dynamically establishing lightpaths in a network in which the number of wavelengths is bounded.

Two central issues are common to the assignment of wavelengths to lightpaths under the various policies. First, since as a consequence of resulting photonic switch sizes and current technology, wavelengths are a scarce resource, it is necessary to establish lightpaths *efficiently* in terms of the total number of wavelengths required. Second, the requirement for establishing a lightpath using the same wavelength throughout its route, introduces a potential bandwidth loss when compared to a lightpath establishment in which the continuity constraint is not imposed. This loss can be perceived either as an increase in the number of wavelengths required to successfully establish a given set of lightpaths, or as an increase in lightpath blocking probability, if the number of wavelengths is limited. In providing *efficient* solutions for lightpath establishment, our objective will be to find algorithms that minimize this loss.

In deriving a lightpath establishment algorithm, we first analyze the complexity of an optimal assignment of lightpaths, introducing the following model. We represent the network by a triplet $G(V, E, W)$ in which V represents the set of N nodes, $N = |V|$, E represents the set of *directional* fiber links between nodes in V, (assuming $(u, v) \in E$ if and only if $(v, u) \in E \ \forall u, v \in V$) and W is the set of wavelengths on each link, $|W| = \omega$. It is assumed that ω is equal for all links. We define a *lightpath request* for connecting a given source/destination node pair by the links constituting a path between them. To establish a lightpath, it is necessary to find an unallocated, identical wavelength, on all the lightpath's links.

Definition : Static Lightpath Establishment (SLE) problem - given a network $G(V, E, W)$, $\omega \geq 3$, and a predefined set of lightpaths L, is it possible to establish all lightpaths in the set ?

By showing that the SLE problem is equivalent to the n-graph-colorability problem it is possible to prove its NP-completeness, as given in Appendix A. That is, finding the minimal number of wavelengths that would accommodate the demands would amount to finding the chromatic number of some (general) graph, where the number of colors, n, corresponds to the number of wavelengths, ω.

5.2 Centralized Solutions

Due to the high complexity of an optimal solution we study heuristics solving the lightpath establishment problems. The fundamental aspects of lightpath communications are covered by considering two possible objective functions : minimization of the required number of wavelengths and minimization of lightpath blocking probability.

Before proceeding to describe solutions to the lightpath establishment problem we develop a lower bound on the number of wavelengths required by an optimal algorithm: the number of wavelengths required to establish a given lightpath set without the wavelength continuity constraint. This number is given exactly by the number of lightpaths passing on the "busiest" link (i.e. the degree of edge congestion) and it is also, a lower bound on ω. We term this lower bound policy Non-Wavelength Continuous (NWC).

5.2.1 Static Demands, Unbounded Number of Wavelengths

The first case to be studied is the one corresponding to the Static Lightpath Establishment problem : a network in which a set of n lightpath requests is predetermined and the objective function is to establish all demands using a minimum number of wavelengths.

We use a greedy allocation heuristic which iteratively allocates a given wavelength to all possible edge disjoint (i.e. non-colliding) lightpaths to whom a wavelength was not yet allocated. The procedure terminates upon allocating a wavelength to each lightpath.

Using an intuition first observed in task scheduling problems [28], we first sort the lightpaths according to their respective lengths, and then try to allocate the wavelengths to the *longest* lightpaths first. Intuitively, a long lightpath is harder to establish, since an unallocated identical wavelength must be found on more links. Therefore, by establishing long lightpaths first, a better wavelengths re-use should be achievable, leading to an overall smaller requirement of wavelengths for a given lightpath set.

The exact description of the solution uses the following data structures :

lpcm[i,j] : the lightpath collision matrix. *lpcm[i,j]* = *1* if lightpaths i and j have a link in common (lightpaths i, j collide)

lpnum[i] : lightpath id's, ordered by descending lightpath length

w : wavelength number currently assigned

set[i] : sets of lightpaths ordered according to allocated wavelength

s, e : start, end pointers to current set

lambda[i] : wavelength definition array. *lambda[i]* points to the first lightpath in *set* using wavelength i

lpalloc[i] : flags indicating if lightpath i was already allocated

n : number of lightpaths in set

or(set,s,e,lpnum[i],lpcm) : function; returns *true* if lightpath *lpnum[i]* has a link in common with the lightpaths in the set *set[s]..set[e]*, based on the lightpath collision matrix *lpcm*.

```
procedure static_establish
begin
  lambda[1] = w = s = e = 1
  for i = 1 to n do lpalloc[i] = false
  while (e < n) do begin (*)
    for i = 1 to n do begin
      if not lpalloc[i] then
      if not or(set,s,e,lpnum[i],lpcm) then begin
        set[e] = lpnum[i]
```

```
            e = e + 1
            lpalloc[i] = true
          end
        end
        w = w + 1
        lambda[w] = s = e;
      end
    end
```

5.2.2 Static Demands, Bounded Number of Wavelengths

Since the number of available wavelengths in WDM systems is expected to remain limited, it is of importance to study the problem of establishing a given set of lightpaths when the number of wavelengths, ω, is bounded by *omega*. We note that in this case, it is possible that new lightpaths will be blocked. The objective function therefore changes in this case to minimizing the ratio of lightpaths rejected to lightpaths requested, defining the lightpath blocking probability.

The previous heuristic maximizes the use of every wavelength before proceeding to allocate a new one. Thus, in effect, it intuitively maximizes the number of unused wavelengths in the network in case their number is bounded. Noting that as long as there is an unused wavelength, the lightpath blocking probability will be zero, we employ a variation of this heuristic for the bounded wavelength problem. We note that an existing drawback of this objective function is that it does not differentiate between long and short lightpaths. Hence, a policy using this objective function, will in effect, discriminate against long lightpaths. The relative effects of blocking probability as a function of lightpath length are studied in section 5.2.5.

As before, we shall allocate a given wavelength to all possible lightpaths that have not yet been allocated a wavelengths. However, the procedure will stop *either* if all lightpaths have been allocated a wavelength or the available wavelength pool has been exhausted. In addition, to allow for an unbiased study of the effect of blocking probability as a function of lightpath length, we avoid sorting lightpaths according to lightpath length

as in the previous case. Thus, the heuristic given in section 5.2.1. remains unchanged, except the line marked by (*) in the algorithm which is modified to :

while $(e < n)$ <u>*and* $(w < omega)$ *do begin*</u>

5.2.3 Dynamic Demands, Unbounded Number of Wavelengths

The interest in studying the case in which lightpaths are established and terminated dynamically, stems from the fact that the Lightnet topology can be modified by reassigning lightpaths. By establishing the lightpaths dynamically, the Lightnet can be reconfigured for purposes of reliability, availability or even adaptation to long term traffic patterns.

We observe that in addition to the efficient use of wavelengths, the issue of *stability* becomes of primary importance in the dynamic case. Past experience with dynamic resource allocation suggests that lightpath allocation solutions might display a "fragmentation" problem in which, while wavelengths may be available on each link on a given path between a source and destination, the continuity constraint over the total path is not satisfiable. Hence it is important to establish whether a given allocation algorithm deteriorates over time as it does for example, in many memory allocation schemes.

We note that the approach developed for the static cases maximizes the use of every wavelength it allocates, before proceeding to allocate a new wavelength. We therefore pursue this approach for establishing lightpaths dynamically, as it intuitively leads to the maximal reuse of wavelengths, or in other words, should reduce fragmentation. The above approach can be mimicked in a dynamic environment by a greedy heuristic that establishes each lightpath using the first available wavelength. Thus, a new wavelength will be allocated if and only if a lightpath cannot be established using any of the wavelengths already in use.

We first consider the case of an unbounded number of wavelengths. In the formal representation for this solution we shall use the following data structures :

lightpath[id] : lightpath information record, holding the following fields :

- *path:* the links constituting the lightpath

- *len:* the lightpath length

- *wavelength:* the wavelength assigned to lightpath

busy[i,j] : *busy[i,j] = 1* if wavelength j is currently assigned to a lightpath passing through link i

path(s,d,vec,len) : procedure; returns in *vec* a route for a lightpath from s to d, of length *len*

getid(id) : function; assigns a unique *id* to a lightpath

wave : index, used in searching for an available wavelength

The establishment procedure scans the matrix *busy* by columns (wavelengths) attempting to find a column where all the entries corresponding to the lightpath's links are zero (unused). If no such column is found among the wavelengths currently in use, the wavelengths counter, *wave*, is increased so as to allocate a new wavelength. Following are the procedures used to establish and terminate lightpaths :

```
establish(s,d,id)
(* establish a lightpath from s to d *)
begin
  path(s,d,vec,len)
  getid(id)
  lightpath[id].path = vec (* save path for hangup *)
  lightpath[id].len = len
  (* find wavelength in which to establish lightpath *)
  found = false
  wave = 1
  while not found do begin
    tmp = 0
    for i = 1 to len do tmp = tmp + busy[q[i],wave]
    if tmp = 0 then found = true
    else wave = wave + 1 (*)
  end
  lightpath[id].wavelength = wave
```

```
  (* update data structure - lightpath established on wavelength wave *)
  for i = 1 to len do busy[q[i],wave] = 1
end
```

Lightpath termination is taken care of by the following procedure :

```
terminate(id)
(* terminate a lightpath *)
begin
  for i = 1 to lightpath[id].len do
    busy[lightpath[id].path[i],lightpath[id].wave] = 0
end
```

5.2.4 Dynamic Demands, Bounded Number of Wavelengths

Last, we consider the problem of dynamically establishing lightpaths in a network in which the number of wavelengths is bounded by *omega*. As before, the problem is motivated by the limitation on the number of wavelengths.

Following the reasoning of the preceding unbounded case, we again employ a greedy approach. The wavelengths are checked in sequential order, establishing a lightpath by allocating it the first wavelength that is not in use on any of the lightpath's links. However, in this case, as the number of wavelengths is bounded, lightpath requests may be blocked. The heuristic, therefore, proceeds as before except in this case, prior to increasing the number of wavelengths in use, it is checked if the maximal limit has been reached. Thus, by changing the line marked by (*) in the previous heuristic to

```
else if wave < omega then begin
    wave = wave + 1
```

the solution for the bounded wavelengths case is obtained.

5.2.5 Results

The performance of the lightpath establishment heuristics is evaluated in terms of the efficiency of wavelength allocation. Having shown that the exact solution is NP-Complete, comparison to exact results is not feasible for any networks of interest. However, as pointed out in section 5.1, by removing the wavelength continuity constraint from lightpath establishment a lower bound on the number of wavelengths needed is obtained. Thus, a comparison to the lower bound obtained by non-wavelength-continuous (NWC) case, can be made to evaluate the performance of the various heuristics as well as to determine the relative penalty imposed by the continuity constraint of the proposed lightpath establishment solutions.

The performance of the presented heuristics was derived by simulating general topology networks under varying traffic conditions and objective functions. All results were obtained with a confidence level of 95%. Lightpaths were randomly created choosing source/destination nodes according to a uniform distribution. The links constituting each lightpath were chosen following a shortest path policy, assuming all links to be of unit length, with random tie breaking rule. For the dynamic environments, lightpath arrival rate refers to the number of lightpath establishment requests per unit of time. An arrival rate λ is implemented in the simulation as an exponentially distributed lightpath request interarrival time with mean $\frac{1}{\lambda}$. Lightpath holding times were assumed to be deterministic, and equal to 200 time units.

In Table 3 we study the case of unbounded number of wavelengths by observing the average number of wavelengths required to establish a given lightpath set size (static demands) for three different network sizes. For each set size, the results presented are averaged over 10 different randomly generated lightpath sets. We observe that the results for the proposed policy and the NWC lower bound are practically identical. This result can be explained by considering the implications of Theorem 2 (see appendix A). The only discrepancy that may arise between NWC and the proposed policy can occur only when cycles are contained in the network graph, with the lightpath demand set also forming a cycle. However, the probability of such a structure occurring, given a lightpath set, is much smaller than the probability of multiple lightpaths passing through a link

in the network. Hence, with high probability, the most congested link in the network determines the number of wavelengths required by the lightpath establishment policy as well as determining the NWC lower bound. The study of the effect of network size on the number of wavelengths required to establish a given demand set, shown in Table 3, supports the above observation. In addition it shows, that as the network size increases, the number of wavelengths required for a given set size decreases. This is due to the fact that in a larger network there are fewer collisions between lightpaths for the same lightpath set size.

In Figure 9 a system with static demands and a topology depicted by Figure 10 where the number of wavelengths, ω is set to 5, is studied. The objective function in this case is the minimization of lightpath blocking probability given in Figure 9 as a function of the lightpath set size. The average blocking probability for the wavelength continuous policies, shown in Figure 9, is *lower* than the NWC lower bound, by up to 2%. This apparent contradiction is explained by observing that the wavelength continuous policy exhibits a 4% higher blocking probability than NWC when considering only long lightpaths (in this case equal to the network's diameter, 5 hops). Since a long lightpath takes up system resources that can be used by multiple short ones, an average lower average blocking probability results when long lightpaths are blocked. Figure 9 also shows the blocking probability for short (1 link) lightpaths confirming the above observation, noting that lightpath blocking probability for the wavelength continuous policy is lower than that of the NWC case.

Before studying the actual number of wavelengths required to accommodate systems with dynamic demands and an unbounded number of wavelengths (ω), or the blocking probabilities in systems with dynamic demands and bounded ω, we first verify the *stability* of these results. Figures 11 and 12 study the stability of the heuristics for the unbounded and bounded wavelengths cases respectively. Figure 11 considers the stability of the proposed heuristics in an unbounded wavelength network by plotting the number of wavelengths required to establish all demands for three different lightpath set sizes as a function of time. It is shown that following the transient phase, the average number of wavelengths does not increase over time. Similarly, Figure 12 shows the blocking proba-

bility as a function of time for the bounded wavelengths case ($\omega = 5$). We note that in this case the blocking probability also remains practically constant over time.

A dynamic establishment of lightpaths without having the ability to perform wavelength reallocation to already established lightpaths, can be expected to have a notable bearing on dynamic lightpath establishment heuristic. Table 4 displays the number of wavelengths required for the lightpath solution and the NWC lower bound as a function of lightpath request rates. Comparing the values corresponding to the lightpath establishment heuristic with the NWC allocation we observe that for high rates, less than 25% additional wavelengths are required on average to establish lightpaths for the same lightpaths request arrival rate. We also observe that the ratio between the number of wavelengths required under wavelength continuous and the NWC case remains almost invariant for different network sizes. Last, when observing the same lightpath request arrival rate over different network sizes, the absolute number of wavelengths required decreases, for reasons similar to the ones stated in the static case.

Figure 13 shows the case were the number of wavelengths is set to 5. The objective function in this case is the minimization of lightpath blocking probability, depicted in these figures as a function of the lightpath request rate, comparing the presented heuristic with the NWC lower bound. We observe that the heuristic performs with a relatively small penalty relative to the optimum.

Finally, it is of interest to investigate the relative improvement obtainable by increasing the number of available wavelengths. Figure 14 depicts the lightpath blocking probability as a function of the number of available wavelengths, ω, for various lightpath request rates in a dynamic lightpath establishment environment. It is noted that for lower rates, a small increase of ω leads to a substantial reduction in blocking probability, whereas high rates require a large increase in the number of wavelengths to obtain a similar blocking probability reduction. We further observe that the blocking probability for a request rate of 0.5 tends to zero for $\omega > 20$. With higher rates the blocking probability increases, reaching 0.47 for the same ω and a rate of 2.0. Thus, we conclude that small increases in the number of available wavelengths can provide substantial reduction in lightpaths blocking probability for small (less than 0.5) lightpath request rates.

5.3 Distributed Solutions

When dealing with an environment where lightpaths are requested and terminated dynamically a distributed solution for efficient lightpath establishment becomes attractive. The performance of a distributed heuristic can be studied from two perspectives. The first, similar to the centralized solution, is concerned with the "performance penalty" of the lightpath approach and the continuity constraint thereof in terms of blocking probabilities. To obtain this, we perform a comparison with the NWC policy where lightpaths are established for as long as there is *any* resource available on each link. The second issue is to compare the results obtained in a distributed way with those that can be obtained if all information is available, in a centralized way, through a Centralized Lightpath Allocation (CLA) heuristic. The "ideal" NWC case was defined in the preceding section, we therefore next define the CLA heuristic.

The CLA solution is based on the following principle of achieving maximal wavelength reuse throughout the network:

As long as there is at least one wavelength λ_i which is not allocated on any link in the network, CLA guarantees that any new lightpath demand will be met with no blocking.

The CLA algorithm thus operates as follows:

Assume that a given wavelength λ_i has already been allocated in a subset $E_i \in E$. The larger this subset, the smaller is the proportion of new lightpaths which can be established allocating λ_i. Thus, for any new lightpath demand that can be established, using one of $\lambda_1, \lambda_2, \ldots$, we should perform this allocation by assigning it the wavelength λ_i (in the group) with largest E_i set.

The above CLA heuristic is evaluated using a simulation with the following parameters : Lightpath duration times were taken as constant (200 time units) and lightpath inter-arrival time as exponential. All results were measured with a confidence level of 99%. Traffic was assumed to be uniformly distributed; routing was non-alternate, shortest path, choosing a path at random when several were possible. 10 wavelengths were assumed to be available on each link. In figure 16, the blocking probability is given as

a function of lightpath arrival rate measured in lightpaths per time unit for the general topology network depicted in figure 15. Figure 16 depicts the blocking probabilities for CLA and "conventional circuit switching" (NWC policy) averaged over all lightpaths and for longest lightpaths only. In terms of average blocking probability we see that the results are very close and in fact, for certain loads, CLA actually displays slightly lower blocking probabilities. This is due to the fact that long lightpaths are rejected by CLA with a higher probability than in conventional circuit switching (see figure 16). The higher rejection probability occurs due to the requirement to find an identical free wavelength throughout the path. Hence since with CLA more short lightpaths will be established, the average blocking probability is decreased. Having defined the ideal case and a corresponding centralized heuristic, we now proceed with defining the distributed solution.

5.3.1 PACK - A Distributed Heuristic

Lacking the global information used in CLA to maximize resource re-use across the network, two viable approaches for distribution can be taken. The first approach is based on exchanging information between neighbors, eventually creating a global picture, or an assessment thereof, in each node. This approach is useful when the lifetime of the information is long with respect to the information propagation time. However, when the structures described have connection times that may be short, nodes will be making decisions based on outdated information most of the time. We further point out that this approach also incurs an additional complexity cost in bandwidth dedicated to control. The second alternative is to emulate global knowledge by implementing a global policy. This can be done by requiring that nodes that decide which resource will be allocated to a lightpath do so, by using the same rules. This is precisely what the following "PACK to beginning" heuristic does. Let $\lambda_1, \lambda_2, \ldots$ be any arbitrary numbering of the wavelengths known to all nodes. PACK will allocate the *smallest* numbered wavelength feasible. Thus on a new lightpath requirement PACK emulates CLA in the attempt to maximize re-use of wavelengths allocating them in the same order in all nodes. However, CLA adapts this order according to the current wavelength allocation while PACK uses a fixed, preset

ordering. Hence, CLA has a superior ability to adapt to changes in lightpath demands. Performance discrepancies between these two heuristics may therefore be expected due to this difference.

In the PACK distributed solution four types of messages are exchanged between the nodes. The message length is, in the worst case, $O(\omega)$. These messages are :

REQUEST (src,dest,wave,id) : lightpath establishment request. *wave* is a bitvector containing a "0" in the i'th location if wavelength i can be allocated for the lightpath. *id* is a unique lightpath identifier, obtained locally by concatenating the originating node id to some counter.

ACCEPT (src,dest,i,id) : lightpath establishment notice. i is the wavelength number allocated to the lightpath.

REJECT (src,dest,id) : rejection notice issued when a lightpath request is blocked.

HANGUP (src,dest,id) : lightpath termination message, initiated by node originating the lightpath request.

Each node maintains the following data structures :

lightpath(id) An array containing a record for each lightpath passing through the node. The record contains the wavelengths allocated (or reserved) for the lightpath, its incoming edge and its outgoing edge.

switch[1..n,1..d,1..d] where $\omega = n$ and d is the degree of the node. **switch[]** defines the wavelength allocation and the appropriate switching function in the node (e.g. $switch[3, 2, 4] = 1$ indicates that link 2 is to be switched to link 4 for wavelength λ_3).

Following is the algorithm executed by each node upon receipt of the corresponding messages :

PACK algorithm

```
request(src,dest,wave,id)
    if dest = node then begin
        i = select_ slot(wave)
        accept(src,dest,i,id)
    end
    else begin
        n = next_ node(dest) (* next node in route *)
        for every wavelength i
        if i used in incoming/outgoing link then
            wave[i] = 1
        if wave[i]=1 ∀i reject(src,dest,id)
        else begin
            update data structures for lightpath id
            if wave[i] set to '1' in this node,
                set to '1' the relevant entry in switch
            send(n,REQUEST,node,dest,wave,id)
        end
    end
end

accept(src,dest,i,id)
    set to '0' all entries previously set to '1'
        in switch for id, except i
    let n be the incoming node of id
    (from lightpath(id))
    if src ≠ node send(n,ACCEPT,src,node,i,id)
end

reject(src,dest,id)
```

```
    set to '0' all entries previously set to '1'
        in switch for id
    let n be the incoming node of id
    (from lightpath(id))
    if src ≠ node send(n,REJECT,src,node,id)
end

terminate(src,dest,id)
    if node ≠ dest begin
        free entry in switch corresponding to id
        let n be the outgoing node of id
    (from lightpath(id))
        send(n,TERMINATE,node,dest,id)
    end
end
```

where *select_ slot* return the lowest numbered feasible wavelength.

Figure 17 contains a comparison in terms of blocking probabilities for PACK and CLA for the sample network. As can be seen, the results are practically identical, both for the average length lightpath blocking probability and the longest lightpath blocking probability. Hence, by transmitting information only along the path of the lightpath, we have obtained, contrary to intuition, a distributed heuristic paying a negligible price in terms of performance.

6. Virtual Topology Construction

As pointed out earlier, the Lightpath architecture constructs a *virtual topology*, the *Lightnet*, in which lightpaths are the new links. Transmissions between any two nodes in the Lightnet take place on the lightpaths, passing on the way photonic switches within the lightpath and electronic switches between lightpaths.

The virtual topology determines the set of lightpaths needed, the passive and active switch sizes, the load balancing properties of the system, routing, congestion, and other control procedures, and determines the hardware requirements and performance. The exact choice of the virtual topologies is therefore clearly a key aspect in the proposed approach. For instance, when using a virtual tree topology, the node degree η can be determined arbitrarily (e.g. $\eta = 3$ for a binary tree), while the total number of wavelengths is bounded by $\frac{1}{2}(\eta - 1)\log_{\eta} n$ for a network with n nodes [24]. On the other hand, a tree topology creates inherent traffic bottlenecks at all subtree roots with a potentially heavy performance degradation for most traffic patterns. In this phase of research we define the embedding process, propose an embedding solution and choose the embedding of a hypercube regular topology which eliminates traffic bottlenecks.

6.1 Problem Definition and Demonstration

In Lightnet, a virtual topology is embedded in a general topology by associating the nodes of the virtual topology with the nodes of the original, general topology and implementing the edges of the virtual topology by creating lightpaths. The embedding process consists of three issues: first, the mapping of nodes in the virtual topology to nodes in the physical topology, determining a list of source/destination nodes in the physical topology that must be connected by lightpaths. Second, the determination of the physical links constituting each of these lightpaths, henceforth referred to as *lightpath routing.* Third, the allocation of wavelengths to lightpaths, so that 1) the same wavelength is allocated to a lightpath throughout its span, henceforth termed *wavelength continuity,* and 2) no allocation conflict occurs, a conflict being defined as the allocation of the same wavelength to two lightpaths passing through the same link. We observe that for a general set of source/destination nodes and a given lightpath routing, the wavelength allocation problem was addressed in the preceding section. In the context of a hypercube topology embedding, we wish to solve the resulting instance of the wavelength allocation problem in a way which provides bounds on the number of wavelengths required, while exploiting the characteristics of regular topologies.

A polynomial time algorithmic procedure embedding a regular topology graph in a

general topology network, while minimizing the number of wavelengths and consequently the complexity of the associated hardware, is not known [Appendix B, 7]. We therefore introduce a two phase heuristic solution: in the first step we obtain a representation of any general topology network in the *simplest regular* form, a string, and then proceed in the second step to embed the regular topologies in this string. This approach is demonstrated next.

Embedding Demonstration

The two phase approach is illustrated through the embedding of a hypercube in the general topology network of figure 18(a). For this network we first find an equivalent representation as a string, possessing the property that any two paths that are edge disjoint on the string are also edge disjoint in the original network. This property ensures that a wavelength allocated to two edge disjoint lightpaths in the string can also be allocated to the corresponding lightpaths in the original topology, without causing an allocation conflict. Figure 18(b) shows such a string representation. We notice that in this case a one to one correspondence exists between the edges of the string and those of the general graph. However, in the general case, a string edge may correspond to any subset of adjacent general graph edges.

We next establish a sequential mapping of hypercube nodes to string nodes, in which node 0 in the hypercube is mapped to node I in the string, node 1 in the hypercube to node II in the string, etc. Following this mapping, we obtain the lightpaths implementing the hypercube edges as shown in figure 18(c). Finally, an allocation conflict free wavelength assignment, using five wavelengths is also shown in this figure, i.e., all lightpaths passing on every physical link are allocated different wavelengths. The hypercube is redrawn in the standard form in figure 18(d), noting that every *edge* in figure 18(d) corresponds to a *lightpath* in figure 18(c). We next present the switching fabric required to support the Lightnet embeddings in every node.

6.2 Hypercube Solution

Several approaches are possible for embedding the string, in step 1, and the Hypercube, in step 2, of the solution. These carry an inherent tradeoff between the conditions the graph

G_p must meet vis à vis the amount of network resources required (number of wavelengths and passive switch size required to embed, in the second step, a regular topology in the generated string). In the following sections we concentrate on a particular solution for each step and show the graph/hardware requirements for it. Alternative solutions are given in Appendix B.

We model a physical network topology as a directed graph $G_p(V_p, E_p)$ where V_p is the set of nodes and E_p the set of edges. Each edge carries ω wavelengths, determined as given in section 2, by the target, regular topology. We assume that if $(u, v) \in E_p, u, v \in V_p$ then also, $(v, u) \in V_p$. Thus, transmissions on the same wavelength can proceed independently in opposite directions. For $G_p(V_p, E_p)$ we seek an *equivalent string representation*, $G_s(V_s, E_s)$ such that :

S1 $V_s = V_p$

S2 Each edge $e \in E_s$ connecting nodes $u, v \in V_s$ corresponds to a subset $\hat{E} \subseteq E_p$ forming a path in G_p from u to v.

S3 Any two paths that are edge disjoint in G_s are also edge disjoint in G_p where the edges are replaced by the corresponding edge subsets.

Conditions **S1-S3** guarantee that any regular topology embedding on the string G_s will also apply to the physical network G_p, using the *same* wavelength allocation. Condition **S3** further guarantees, that two lightpaths allocated the same wavelength in G_s can also be allocated the same wavelength in G_p. Therefore, bounds computed on the number of wavelengths and the associated node/switch capabilities needed in G_s, will also hold in G_p.

The most attractive approach in terms of hardware is obtained by generating the string through the identification of a Hamiltonian path. We observe that the process of generating a string from a Hamiltonian path is immediate as is the proof of the preservation of conditions **S1 - S3**. In such a path each node has a degree of 2, so that the size of the photonic switches in each node is minimal, with a total of ω 2×2 switches required, one for each of the ω wavelengths in the network.

If the physical layout of the original topology can be controlled, a Hamiltonian path can be established easily. Alternatively, it is known that the problem of finding a Hamiltonian path in a given arbitrary graph is NP-Complete, requiring therefore the use of heuristic solutions [8]. The solution presented in [8] presents an average polynomial time algorithm which finds a Hamiltonian path or establishes that none exists. For graphs created randomly, with a fixed probability of an edge existing between any two nodes, the algorithm was shown to find a Hamiltonian path if such exists with an average time of $O(|V|^3)$.

6.2.1 An Embedding Algorithm

We present the hypercube embedding algorithm and then investigate the number of wavelengths required, the virtual network diameter, and the resulting switch size.

Embedding : Let $G_h(V_h, E_h)$ denote a hypercube with a node $v \in V_h$ numbered by an index i, $i = 0..n = 2^k - 1$, k integer and $n = |V_h|$. Number the nodes in a string G_s from left to right by a single index i, $i = 0..n-1$. Define the identity embedding function by :

$$\mathcal{E}(i) = i, \quad i = 0..n-1 \tag{1}$$

Wavelengths allocation : Scan string from left to right. Define an ordering on the lightpaths based on their left end-node. Lightpath l_i is said to be smaller than lightpath l_j if it's left end-node is to the left of lightpath l_j's left end-node. The algorithm keeps track of the usage of wavelengths per link (a wavelength is used in a link if it was allocated to a lightpath passing through the link) and allocates wavelengths to the lightpaths based on the above ordering. Formally :

procedure alloc()
(allocate wavelengths to lightpaths embedding*
*a hypercube in a string *)*
($w = \frac{2}{3}n$: number of wavelengths required *)*
begin
    for i := 1 to w do

for j := 0 to n-1 do used[i,j] := false

(used[i, j] = true if wavelength i was allocated to a lightpath passing link j *)*

for i := 0 to n-1 do

for all lightpaths with origin i do begin

find a wavelength λ *for which*

used[λ,*k*]=*false* $\forall\ i \le k < d$

(d - lightpath destination *)*

allocate λ *to the lightpath and mark*

used[λ,*k*]=*true* $\forall\ i \le k < d$

end

end

Figure 18(d) illustrates the embedding of a hypercube using as the original topology the graph of figure 18(a). By observing the edges encompassed by each lightpath, as determined by figure 18(d), it is seen that the allocation is indeed conflict free, i.e., lightpaths having links in common have been allocated distinct wavelengths. As can be seen, a total of 5 wavelengths is required.

The properties of the hypercube embedding are as follows : a) a node degree which is logarithmic in the number of nodes, b) an average number of hops bounded by $\log_2 n$, and c) a linear number of wavelengths, given by $\frac{2}{3}n$. The first two properties are inherent to the topology chosen. In the remainder we establish the third property and determine its optimality.

Properties:

1. The maximum number of wavelengths required when embedding an n node hypercube in a string with the identity embedding function and the above wavelength allocation algorithm is given by $\frac{2}{3}n$.

2. The embedding presented for the hypercube is optimal up to a constant factor.

3. The $\frac{1}{2}$-width of the n-node hypercube is given by $w(G_h) = \frac{1}{2}n$.

The preceding properties of the hypercube embedding are formally proven in Appendix B.

6.3 Performance of a Hypercube Based Lightnet

To establish the efficiency of the Lightnet architecture we proceed to compare the performance observed using conventional, store and forward wide area network operation with the performance of the same networks employing the hypercube Lightnet embedding. In this embedding the number of wavelengths required is determined by theorem 1 (see appendix B). In the embedding process the lightpath routing was restricted to the string, i.e. only the string edges are used to construct the lightpaths. The performance measures studied are network capacity, the maximum network throughput that can be sustained while maintaining ergodicity in all queues and average buffering requirements at the maximally loaded node.

In the network model we assume: Packet processing takes place at all nodes on a packet's path in a conventional network and nodes performing switching between lightpaths in the Lightnet. A minimum-hop shortest path routing (in terms of nodes performing switching) with random selection for tie breaking rule is used. Since packet processing in high speed networks is significant, the node capacity is modeled as finite. A packet arriving at a node may enter a processing server if one is available, or join a common queue if all servers are busy. Considering the large discrepancy between the optical transmission bandwidth and the processing and propagation times, the former is considered negligible in the model.

To evaluate the network performance a simulation was developed. In it, packet arrivals are assumed to follow a Poisson distribution and source and destination selection for each packet follows a uniform distribution. Node processing capability was modeled by 3 parallel (packet) servers, each with service rate of 0.1 packet/unit time. Upon terminating service the packet proceeds to the next node. A propagation delay of 100 time units is assumed between physically adjacent nodes. All simulation results were obtained with 97% confidence levels.

The hypercube topology was embedded in three randomly created physical topologies: The first "*homogeneous*" topology was generated with a diameter and node degree distribution matching those of an earlier 1978 Arpanet topology. Two other topologies were introduced to study the relative effects of various bottlenecks on the performance of the store and forward and Lightnet architectures. The "*two-lobe*" topology consists of two clusters of 31 nodes each, randomly connected using the same parameters as the homogeneous network. To this configuration two nodes were added, each with one link to one of the nodes in each cluster. The third, "*elongated*" topology, was generated with a longer diameter of 20 as compared to 15 for the two lobe topology and 10 for the homogeneous topology.

Since the primary issue raised in the design of optical networks is the discrepancy between transmission which in turn leads to bottlenecks and reduced user available throughput, we concentrate on comparing network capacities. Figure 19 presents network throughput as a function of system load for the three physical topologies. The results show that in the random topology network capacity was nearly tripled when using the hypercube embedding. For the two-lobe topology, the hypercube embedding provided a capacity increase by a factor of 4.5. Last, in the elongated topology, network capacity increased by a factor of 7.5 for the hypercube embedding. The superior performance of the Lightnet embedding is due to the reduced number of active switching stages per packet transmission and the inherent load balancing when compared to the conventional network operation. Notice that the throughput of Lightnet is independent of the physical topology carrying the embedding. Although, for the same regular topology, packets will traverse different physical paths in different underlying physical topologies, the data paths in terms of nodes performing switching between lightpaths, the factor affecting capacity, remain invariable. Figure 20 shows the buffering requirements, demonstrating that with the Lightnet approach the increased network throughput does not require additional buffers.

7. Implementation Considerations and Switching Node Design

To employ WDM to implement a Lightnet architecture based on lightpaths, a number of hardware related issues must be examined. Among them, the multiplexing technique and the associated number of wavelengths made available, the devices determining the lightpath span and the switching node design.

7.1 Wavelength Budget

The establishment of lightpaths for regular virtual topologies requires a number of wavelength which is a function of the network size. With wavelength division multiplexing a limited number of high speed channels are available on each link. Current experimental systems are able to carry up to 20 channels, each modulated at 2 Gbit/s [12]. Similar devices approaching 60 channels are considered feasible in the near future [12,13].

7.2 Lightpath Span

An important aspect of the lightpath architecture is the lightpath span. If the lightpath span is large, in particular in the number of nodes it can passively traverse, then the virtual topology can be established in a way which optimizes the number of required wavelengths and the system capacity. Conversely, the choice of the virtual topology might be constrained by the maximum lightpath span.

Three basic scenarios are being considered for addressing the lightpath span issue in Lightnet: 1) electronic regeneration, 2) optical amplification and 3) optical regeneration. In the following subsections we describe these options.

7.2.1 Electronic Regeneration

The first scenario is the simplest technological option using off-the-shelf components. It involves optical/electrical (O/E) and E/O conversion in conjunction with the digital regeneration of each lightpath at every node. Transmission rates of 2.4 Gbs are inherent

to existing systems. Therefore this solution can support lightpaths of several Gbs each. While this is a hardware intensive solution, it provides an optimal S/N ratio of the optical signal, since it never has to travel more than one "hop" before being detected and regenerated. With such a solution, a lightpath can be spanned over an unlimited number of nodes. Put differently, lightpaths can be determined by the requirements of virtual topology embedding, as described in the preceding section, and thus be used to optimize network performance. Conversely, it can be argued that this is wasteful, since many links will have needlessly large S/N ratios. The main disadvantage here lies in the replication of the high-bandwidth electronic clock recovery circuitry needed to regenerate the optical signal.

7.2.2 Optical Amplification

The second scenario is a non immediate, while near term technological option, involving the nonregenerative amplification of the optical signal at each node using optical amplifiers at every hop. The main concern with this solution is the preservation of an acceptable S/N ratio so that the optical signal can be successfully detected at the end of the lightpath. Since the signal's timing information is not regenerated as in the first scenario, phase jitter and amplitude distortion will accumulate within the optical signal as it passes from node to node, being amplified at each node. Care must be taken to ensure that the accumulated noise on the optical signal does not cause an unacceptable error rate in the detection circuit at the end of the lightpath.

Experiments conducted recently have shown very successful use of optical amplifiers [15-20]. For instance, in [15], 25 optical amplifiers were used in series (an amplifier every 80km) to provide transmission of a 2.5 Gbit/s optical signal over 2,223 km of single mode fiber, with a power penalty due to accumulated noise of only 4.2dB. This enormous bandwidth-distance product indicates that the noise accumulation due to the optical amplifiers should not be a problem over moderate distance lightpaths ($<$ 1000km).

The use of optical amplifiers in this way eliminates most of the digital regeneration hardware that would be required in the first scenario. Furthermore, optical amplifiers such as those used in the references have recently entered the commercial market at

approximately $15K$ each [21]. It is expected that this price will drop significantly as the technology matures and competition grows.

7.2.3 Optical Regeneration

The use of optical regenerators is a long term solution, possibly 5-10 years away. The existence of non-linear refractive index within a semiconductor laser amplifier has led to proposals for new signal processing devices. The possibility for performing all optical signal regeneration has been demonstrated [26]. This process can be achieved with coherent light by using the bistable characteristics of a Fabry-Perot laser amplifier. The principle has been demonstrated at 140 Mbit/s. It is now necessary to develop techniques to increase the operating speed and to perform clock recovery. When this technology will mature, this option will be the preferred one in terms of system reliability and bandwidth.

7.3 Switch Design

The Lightnet switch is composed of three main switching components, the passive switch, the lightpath terminating switch (LTS), and the active switch, as shown in figure 2. The passive switch provides the "intra-lightpath" switching for the lightpaths passing through the node and switches the lightpaths terminating at the node to the active switch via the LTS. The active switch connects the local node to the network and performs "inter-lightpath" switching according to the destinations of the packets arriving on the lightpaths initiating/terminating at the local node. Lightpaths initiating at the local node proceed from the active switch, through the LTS, to the passive switch. As discussed in section 7.2 signals departing the passive switch can be regenerated and subsequently multiplexed according to the output links thus removing the restriction on lightpath span.

The design of the Lightnet switch, as given in section 3, requires the use of passive and active switching elements. The first class consists of the passive devices, also referred to as "relational" devices (used for intro-lightpath switching of data). The second class of devices will be referred to as the active devices or "logic" devices (used for inter-lightpath switching). The passive devices perform the function of establishing a large bandwidth

"relation", or a mapping between the inputs and the outputs. The relation is a function of the control signals to the device and is *independent* of the data inputs. Thus, the strength of relational devices is that they do not sense the presence of individual bits passing through them, they only pass them. Due to this bandwidth transparency the fabric bandwidth will be the transmission bandwidth. In the active devices the data that is incident on the device controls the state of the device. The high speed reconfiguration (setup time) requirement for these devices will limit the bit rates of signals that can eventually flow through their fabrics to substantially less than those that can pass through the passive switches fabrics. Thus, the strength of active devices is the added flexibility that results from their ability to sense the bits that are passing through them; while their weakness is that through sensing the bits that are passing through them, the maximum bit rate they can handle becomes limited.

We next present the passive and active switches required for our implementation.

7.3.1 Passive Switches

Recall that the requirements from the passive switches in the Lightnet solution are: 1) very high bandwidth (same as the optical transmission bandwidth), 2) slow reconfiguration rate, 3) low crosstalk, 4) low attenuation and 5) small dimensions.

Mechanical optical switches of dimension up to 40x40, can switch single-mode signals with crosstalk of -90dB and attenuation of around 2dB, thus obtaining crosstalk and attenuation characteristics far better than those offered by electro-optic devices [14]. The use of mechanical switches, despite these characteristics, was not previously considered for data switching networks due to their slow set up speeds. The principle of using preset lightpaths changes this situation dramatically: while the setup time of the mechanical switches is relatively large (50ms) [14], this does not constitute a problem in Lightnet, as lightpaths are not established on a per packet basis and can have lifetimes measured in hours or days. Furthermore, the design of the photonic switch required for lightpaths at the intermediate nodes can benefit from the wavelength continuity property of the lightpaths. We observe that since a lightpath maintains the same wavelength throughout its span, a channel incoming on one wavelength need not be switched to another wave-

length. Consequently, in realizing the photonic switch, it is possible to *group* the channels according to wavelengths prior to switching. The photonic switch can thus consist of ω switching matrices, one for each wavelength. Each of these switches has dimension of $(D_p + D_v) \times (D_p + D_v)$, D_v being the physical node degree and D_v the virtual topology node degree, as contrasted with a substantially more complex, $(\omega D_p + D_n) \times (\omega D_p + D_n)$ switch, that would be required without wavelength continuity. We also observe that the wavelength continuity of a lightpath ensures that no wavelength translation will be required within a lightpath.

7.3.2 Active Switches

The dimension of the active switch required for the Lightnet architecture is $(D_v) \times (D_v)$, as contrasted with a substantially more complex, $(\omega D_v) \times (\omega D_v)$ switch, that would be required for a conventional network architecture with ω wavelengths. For active devices the critical parameter that will determine the switch bandwidth will be its reconfiguration, or switching, time. The only way to create a high speed active switch is using the WDM dimension. For reasons amply demonstrated prior to the deployment of WDM for high speed transmission systems, WDM is the only promising technology to provide the Gigabit throughputs through its inherent parallelism. A switch design in WDM optical technology is a significant departure from the conventional electronic packet switch designs. The first WDM designed switch the HYPASS appeared only recently in the literature [25]. HYPASS is a high-performance packet switch exemplifying a packet-switching fabric using WDMA. In this fabric, the packetized information enters the fabric from the left, where it is initially stored in a First In First Out (FIFO). The objective is to modulate the tunable laser, tuned to the fixed wavelength of the designated output port; pass the information through the transport star coupler; and then receive the information at the desired output port. Prior to accessing the star coupler, it is necessary to check to see if the desired output port is busy. This is accomplished through the specialized control hardware. If an output port is available, the protocol processor associated with the fixed wavelength receivers will turn on the laser associated with the particular output port, allowing light to enter the control star coupler. The tunable receivers attached to the control star coupler can

tune to the wavelength of any of the output channels. If the signal is present, it will signal the input channel decoder to tune the laser to the appropriate wavelength and then command the FIFO to send the current packet to the desired output channel. Note that in this fabric the packet address is converted to the specific wavelength of the output channel. Thus, the address in the fabric is the wavelength of light entering the transport star coupler. While the the switch interconnection topology is new compared to the conventional electronic multistage switches, it uses traditional control protocols designed for multiaccess low speed networks. This choice severely restricts the switch performance. It is currently a widely accepted claim that the architecture as well as the switch control need to be compatible with the high speed WDM technological (components) and physical (end-to-end propagation delay) constraints.

It is true in general, and clearly remains true in our case, that the use of the same technology in the (Lightnet) transmission domain (transmitters, receivers, optical switches, etc.) and in the Lightnet electronic switching domain will guarantee that any ongoing technological developments in WDM components will be utilizable and similarly affect both network aspects. This observation further strengthens the argument for utilizing WDM technologies in the switches as well as in the network.

8. References

1. I. Chlamtac and W. R. Franta, "Rational, Directions and Issues Surrounding High Speed Computer Networks," *Proceedings of the IEEE*, January 1990.

2. R. Linke and A. H. Gnauck, "High-Capacity Coherent Lightwave Systems," *IEEE Jour. on Lightwave Tech.*, Vol. 6, No. 11, November 1988.

3. K. Nosu and K. Iwashita, "A Consideration of Factors Affecting Future Coherent Lightwave Communication Systems," *IEEE Jour. on Lightwave Tech.,*, Vol. 6, No. 5, May 1988.

4. E. Nussbaum, "Communication Network Needs and Technologies - A Place for Photonic Switching?" *Jour. Sel. Areas in Comm.*, Vol. 6, No. 7, August 1988.

5. J. Coudreuse, M. Servel, "Prelude: An Asynchronous Time Division Switched Network", *Proc. 1987 IEEE Int. Conf. on Comm.*, 1987.

6. J. Turner, "Design of an Integrated Services Packet Network", *IEEE Journal on Sel. Areas in Comm.*, Vol. 4, No. 8, November 1986.

7. A. Rosenberg, Personal Communication, February, 1990.

8. M. R. Garey and D. S. Johnson "Computers and Intractability: A Guide to the Theory of NP - Completeness," *W. H. Freeman & Company*, 1979.

9. J. R. Goodman and C. H. Sequin, "Hypertree: A Multiprocessor Interconnection Topology," *IEEE Trans. on Computers*, Vol. C-30, No. 12, December 1981.

10. T. G. Robertazzi, "Toroidal Networks," *IEEE Comm. Magazine*, Vol. 26, No. 6, June 1988.

11. D. K. Pradhan and S. M. Reddy, "A Fault Tolerant Communication Architecture for Distributed Systems," *IEEE Trans. on Computers*, Vol. C-31, No. 9, September 1982.

12. S. S. Wagner, H. Kobrinski, T. J. Robe, H. L. Lemberg, L. S. Smoot, "Experimental Demonstration of a Passive Optical Subscriber Loop Architecture", *Electronics Letters*, Vol. 24, No. 6, March 1986.

13. DiCon Fiberoptics Inc., 950-C Gilman Street, Berkeley, CA 94710.

14. R. Alferness, "Waveguide Electrooptic Switch Arrays", *IEEE Journal on Selected Areas in Communication*, August 1988.

15. S. Saito, T. Imai, T. Sugie, N. Ohkawa, Y. Ichihashi, T. Ito, (NTT Transmission Systems Laboratories, Japan), "An Over 2,200km Coherent Transmission Experiment at 2.5 Gbit/s Using Erbium-doped-fiber Amplifiers", *OFC'90 Postdeadline Papers*, San Francicso, January 1990.

16. N. Edagawa, Y. Yoshida, H. Taga, S. Yamamoto, H. Wakabayashi, (KDD RD Laboratories, Japan), "12300ps/nm, 2.4 Gbit/s Optical Fiber Transmission Experiment using Er-doped Fiber Amplifiers and Penaly due to Transmitter Phase Noise, *OFC'90 Postdeadline Papers*, San Francicso, January 1990.

17. H. Taga, Y. Yoshida, N. Edagawa, S. Yamamoto, H. Wakabayashi, (KDD Meguro RD Laboratories, Japan), "459km, 2.4 Gbit/s 4 Wavelength Multiplexing Optical Fiber Transmission Experiment using 6 Er-doped Fiber Amplifiers, *OFC'90 Postdeadline Papers*, San Francicso, January 1990.

18. W. I. Way, S. S. Wagner, M. M. Choy, C. Lin, R. C. Menendez, H. Tohme, A. Yi-Yan, A. C. Von Lehman, R. E. Spicer, M. Andrejco, M. A. Saifi, H. L. Lehmberg, (Bellcore), "Distribution of 100 FM-TV Channels and Six 622 Mb/s Channels to 4096 Terminals Using High Density WDM and a Broadband In-Line Erbium-Doped Fiber Amplifier", *OFC'90 Postdeadline Papers*, San Francicso, January 1990.

19. K. Kikushima, E. Yoneda, K. Aoyama, (NTT Transmission Systems Laboratories, Japan), "6-Stage Cascade Erbium Fiber Amplifiers for Analog AM- and FM-FDM Video Distribution Systems, *OFC'90 Postdeadline Papers*, San Francicso, January 1990.

20. R. A. Spanke, "Architectures for Guided Wave Optical Space Switching Systems", *IEEE Communications Magazine*, Vol. 25, No. 5, May 1987.

21. BT+D Technologies, Wilmington, Delaware.

22. S. S. Wagner, et. al., "Experimental Demonstration of a Passive Optical Subscriber Loop Architecture," *Electronics Letters*, Vol. 24, No. 6, March 1986.

23. G. R. Hill, British Telecom,Personal Communication.

24. I. Chlamtac, A. Ganz and G. Karmi, "Tree Based Lightnet Topologies," *Tech. Report # CSE-90-10, Dept. of ECE*, Univ. of Massachusetts, February 1990.

25. M. S. Goodman et al., "Demonstration of Fast Wavelength Tuning for a High Performance Packet Switch, *Fourteenth European Conference on Optical Communications*, Brighton, England, Sepetember 1988.

26. R. Alferness, "Waveguide Electrooptic Switch Arrays," *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 7, August 1988.

27. DiCon Fiberoptics Inc., 950-C Gilman Street, Berkeley, CA 94710.

28. R. W. Conway, W. L. Maxwell and L. W. Miller, "Theory of Scheduling," Addison-Wesley, 1967.

Figure Captions

Table Captions

Table 1 : Performance Measures for Lightnet With 128 Nodes

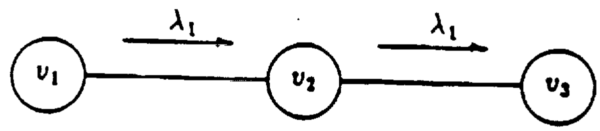
Table 2 : Capacity for Various Network Sizes

Table 3 : Effects of Network Size on the Number of Wavelengths Required for the Static Unbounded Case

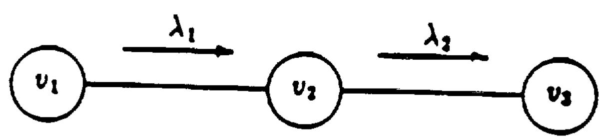
- (a) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 15
- (b) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 30
- (c) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 45

Table 4 : Effects of Network Size on the Number of Wavelengths Required for the Dynamic Unbounded Case

- (a) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 15
- (b) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 30
- (c) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 45



(a)



(b)

Figure 1 : Examples of Lightpath Allocation

$$L_1 : v_1 \rightarrow v_2$$
$$L_2 : v_2 \rightarrow v_3$$

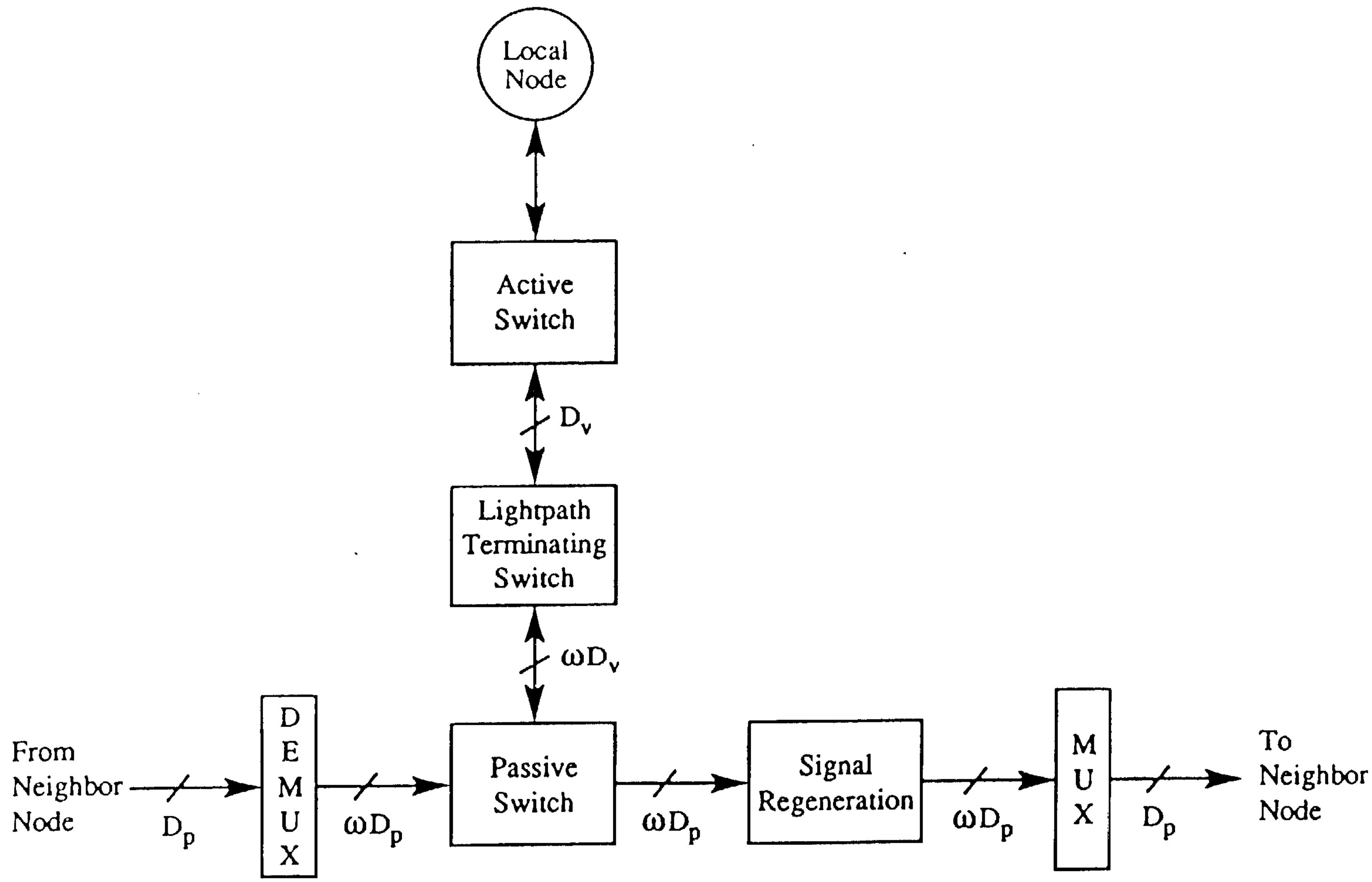


Figure 2 : Lightnet Switch: High Level Design

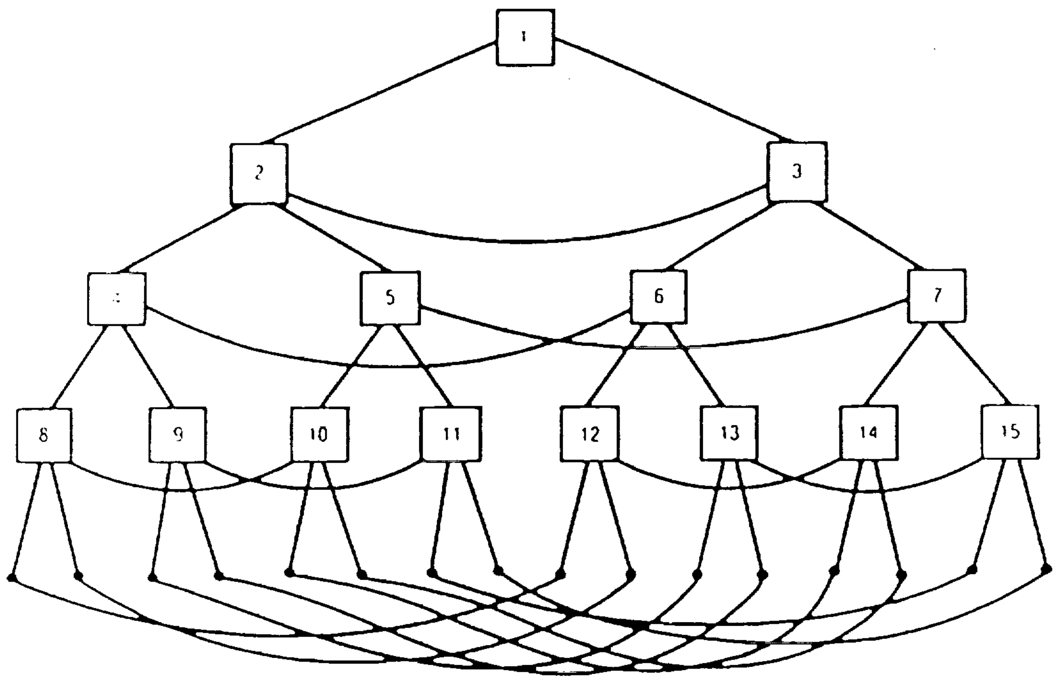


Figure 3 : A Hypertree Topology

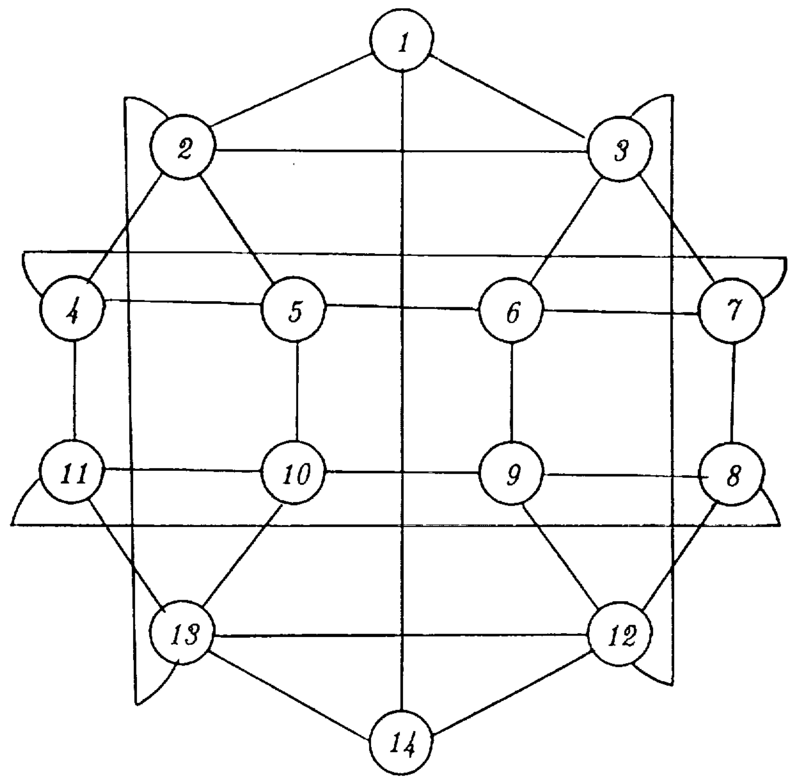


Figure 4 : A Diamond Topology

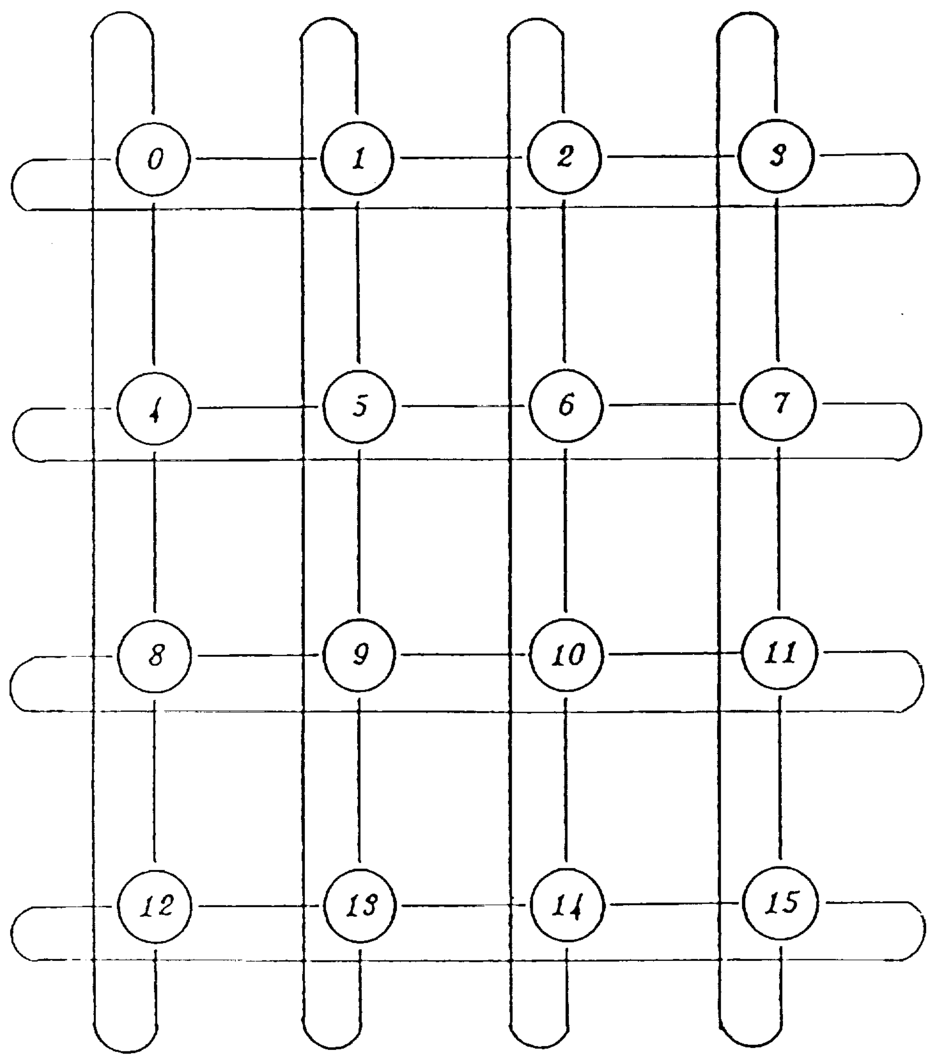


Figure 5 : A Torus Topology

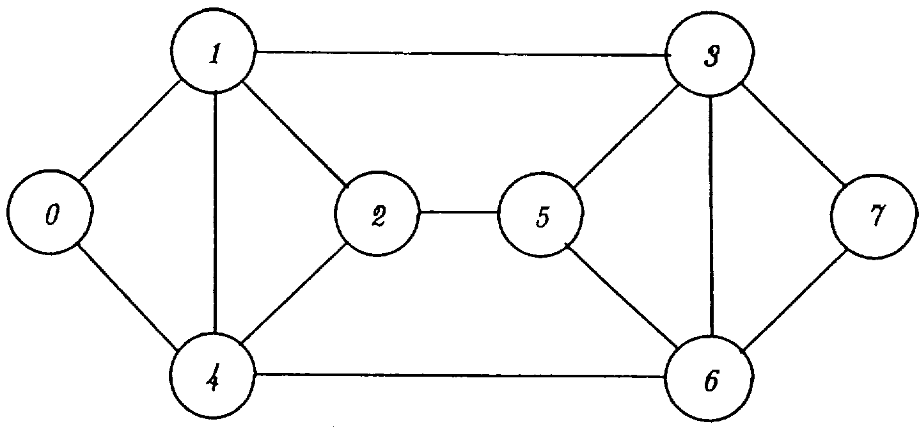


Figure 6 : A De-Brujin Topology

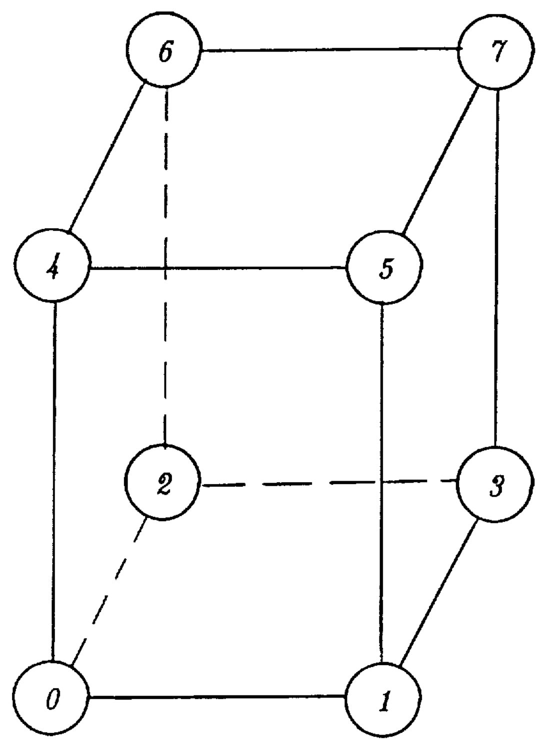


Figure 7 : A Hypercube Topology

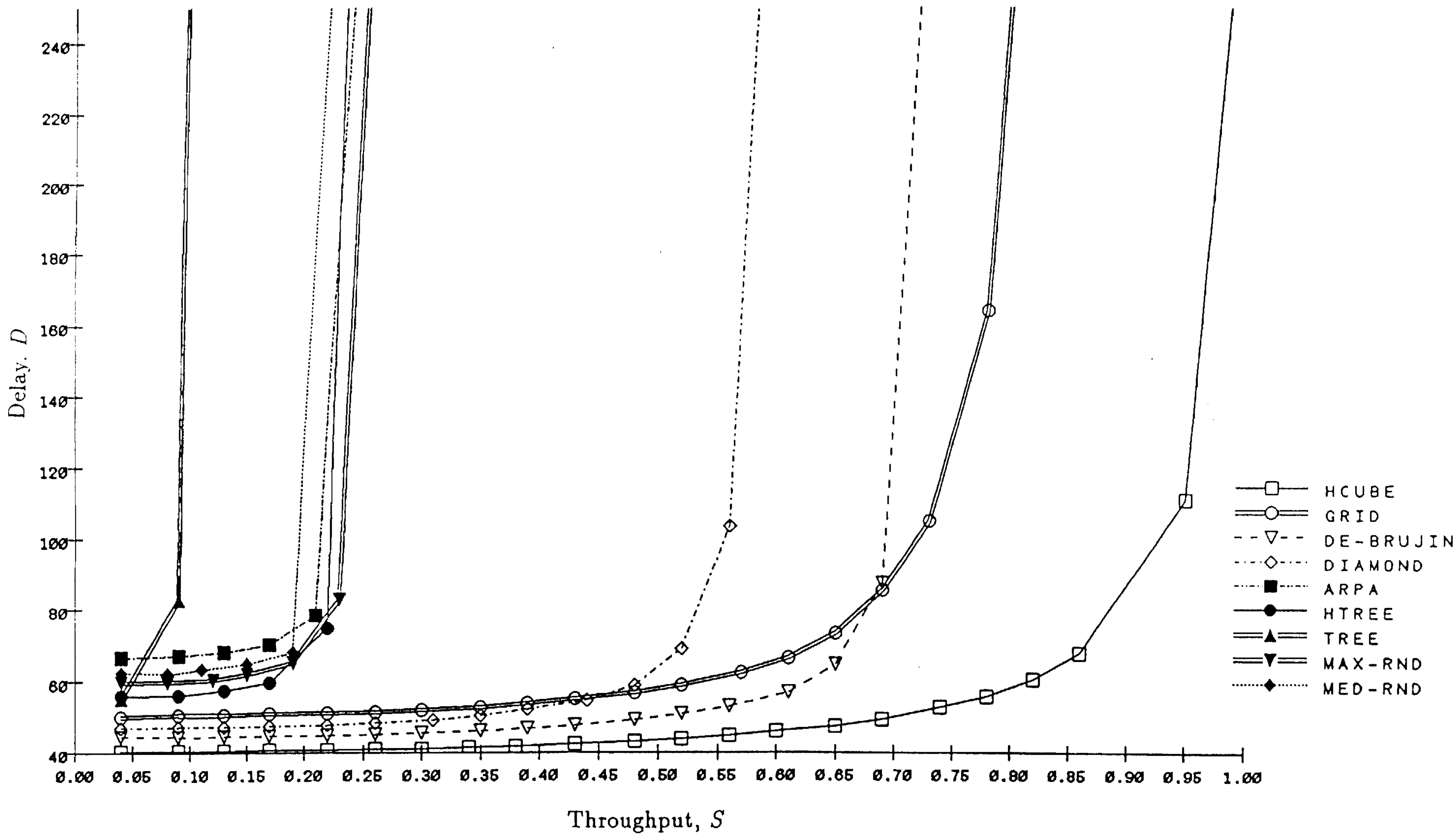


Figure 8: Delay vs. Throughput for 60–64 node networks

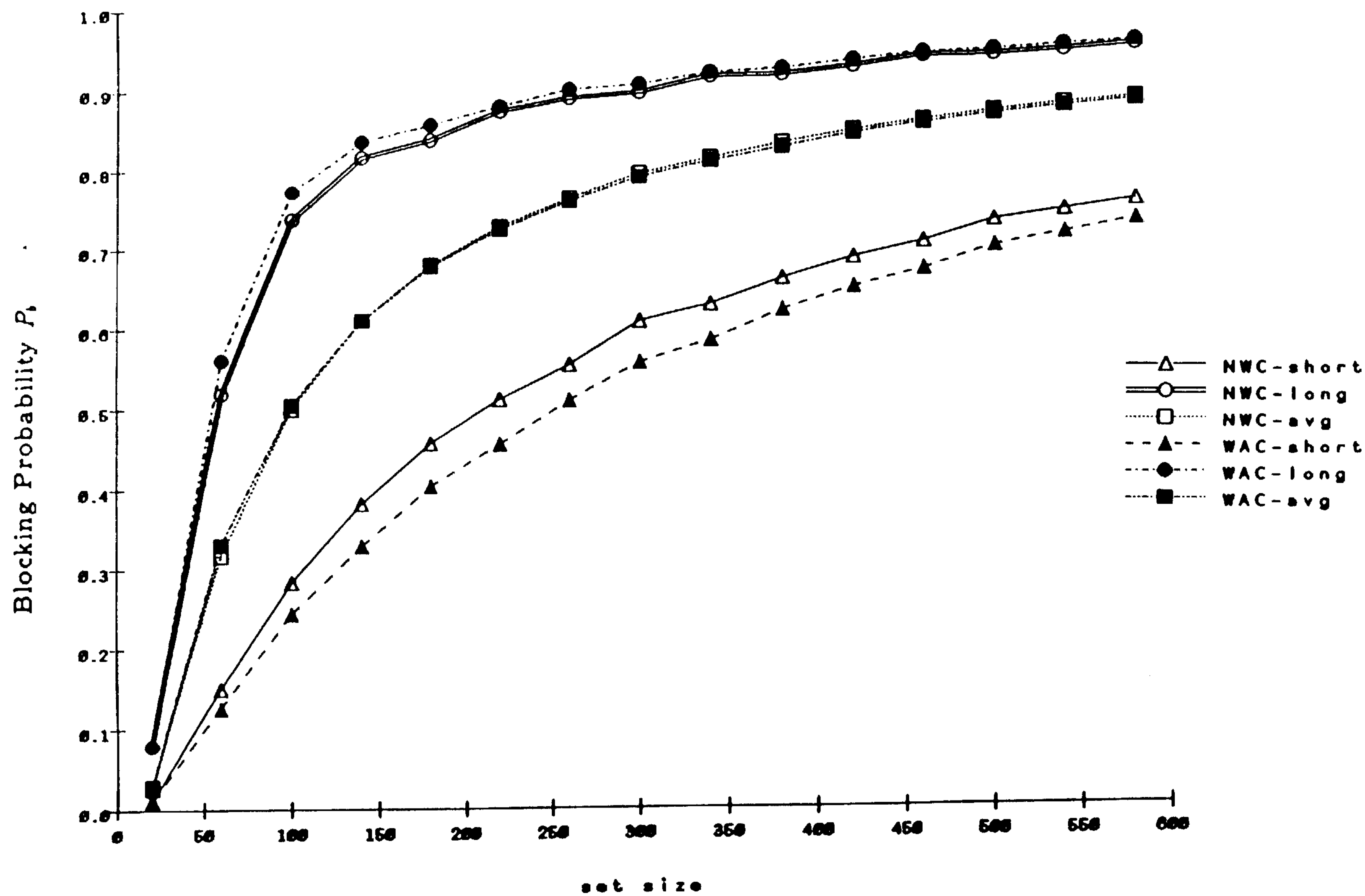


Figure 9 : Blocking Probability vs. Set Size (Static, $\omega = 5$)

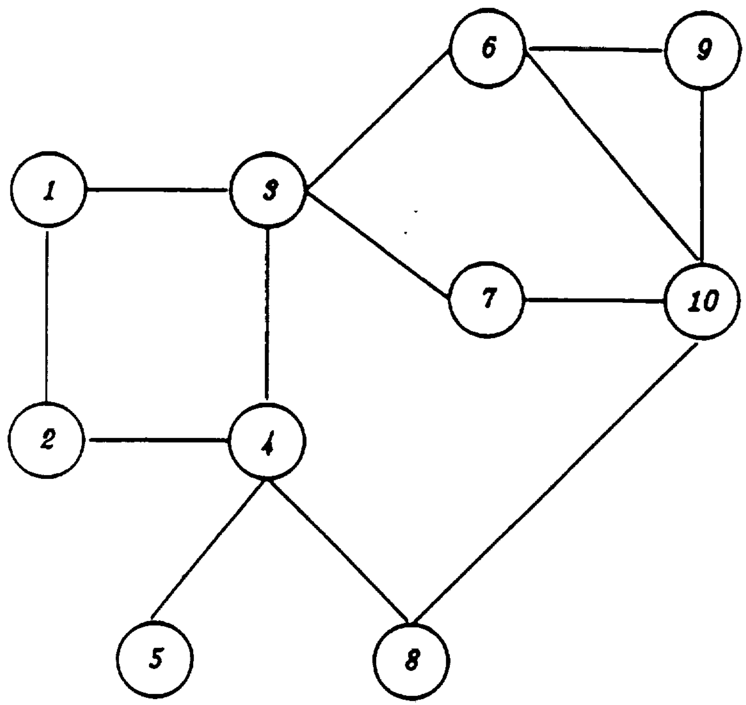


Figure 10 : Sample Network Topology

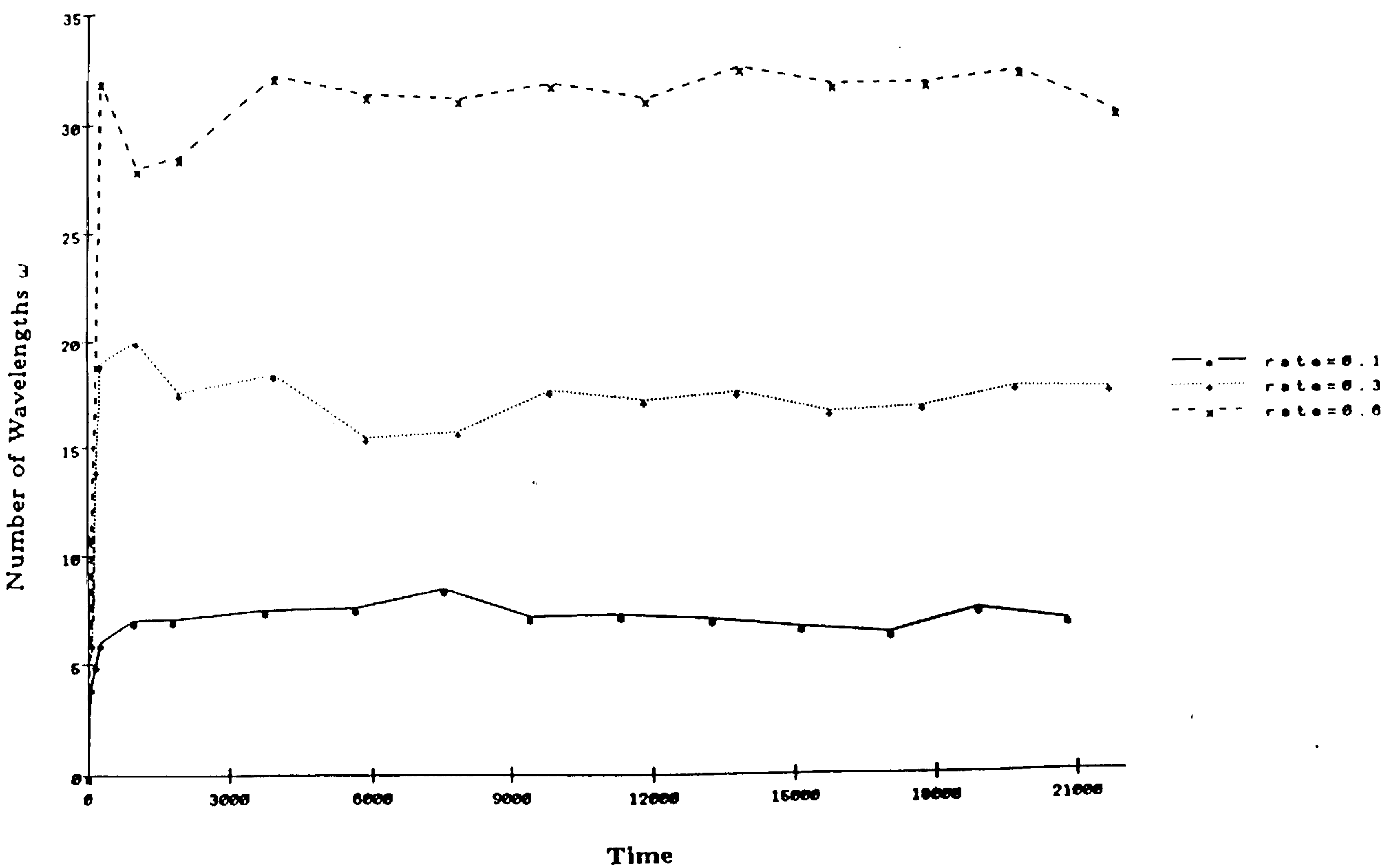


Figure II: Number of Wavelengths vs. Time (Dynamic, Unbounded)

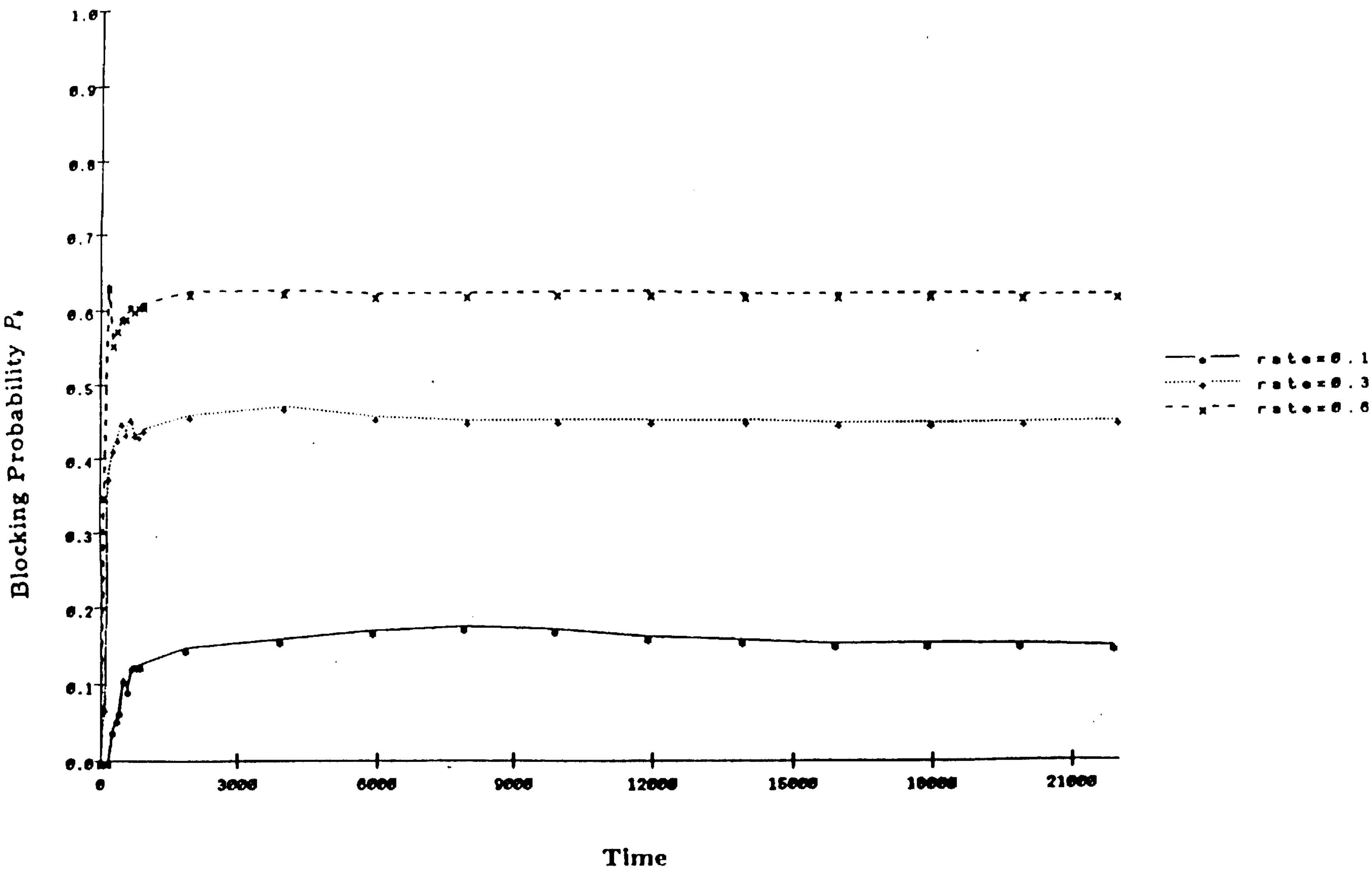


Figure 12: Blocking Probability vs. Time (Dynamic, $\omega = 5$)

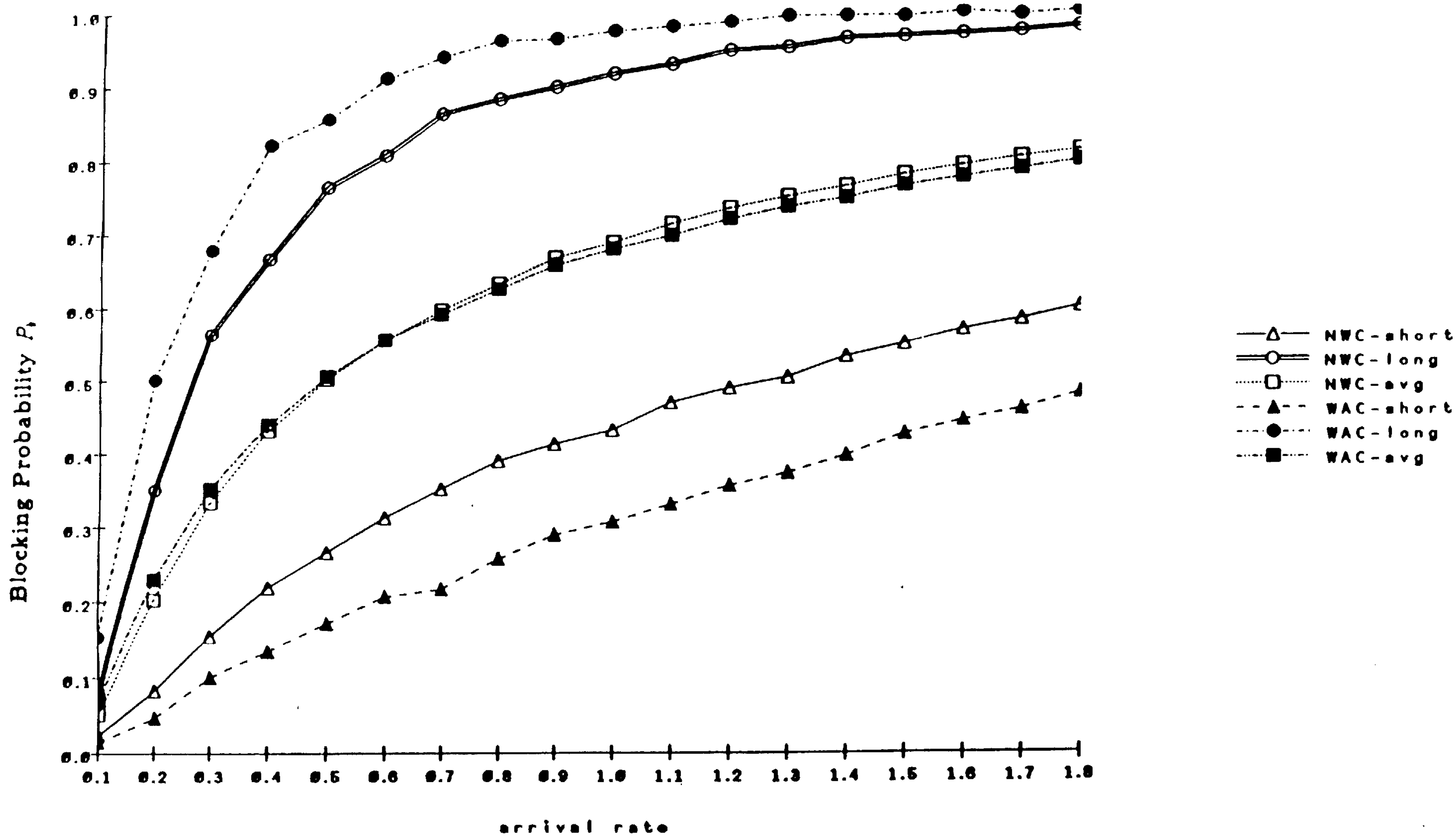


Figure E3: Blocking Probability vs. Rate (Dynamic, $\omega = 5$)

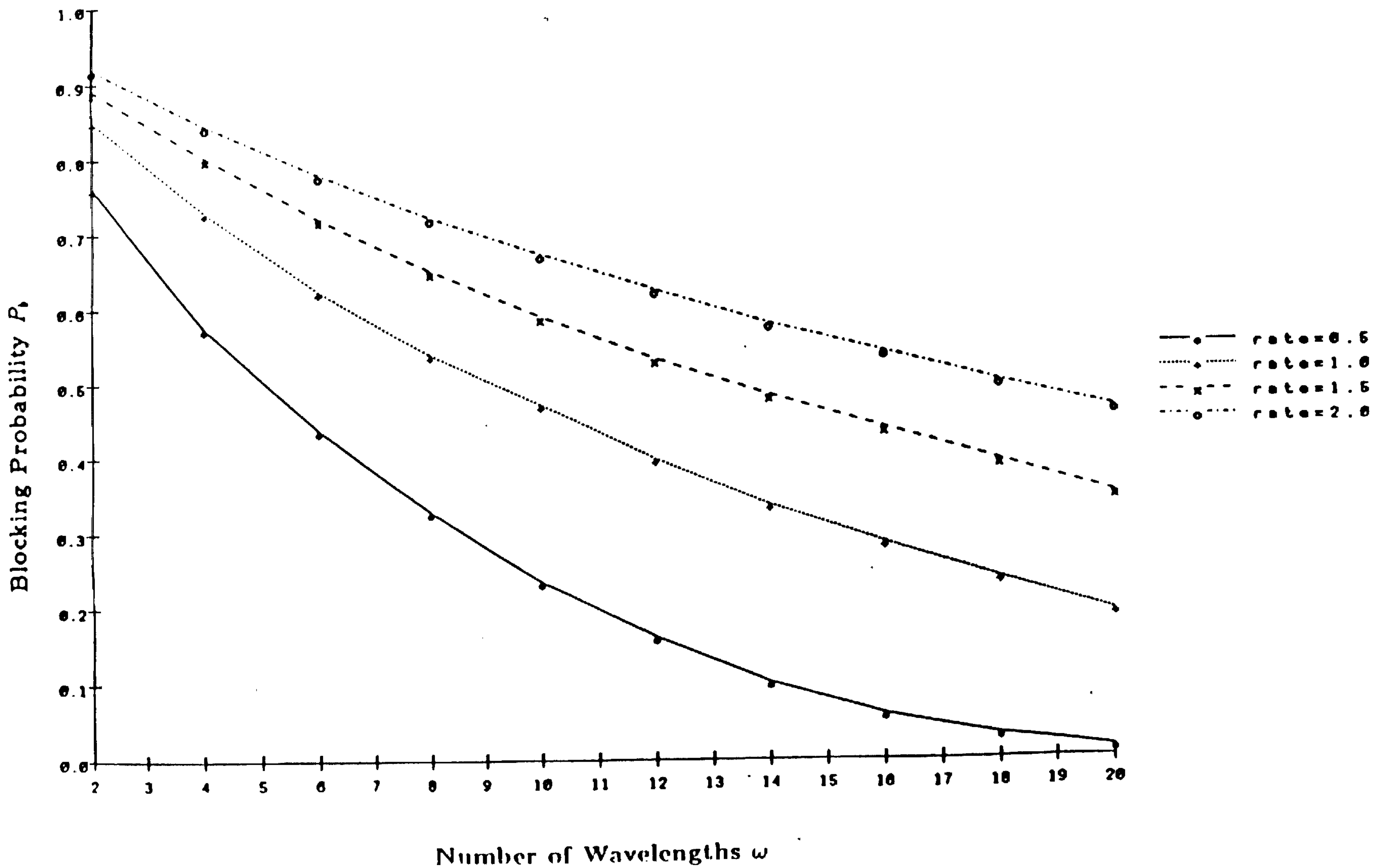


Figure 14: Blocking Probability vs. Number of Wavelengths Bounded, Dynamic

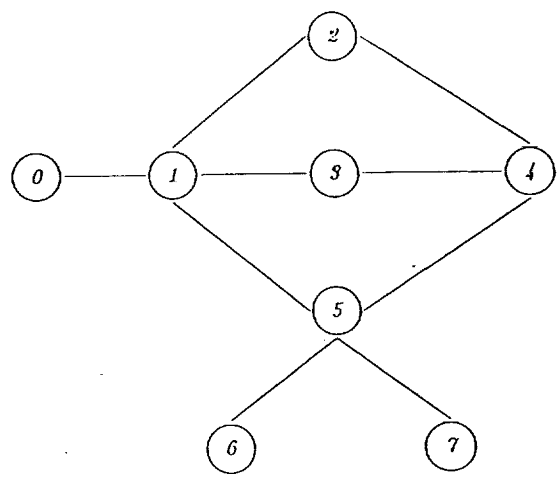


Figure 15: General Topology

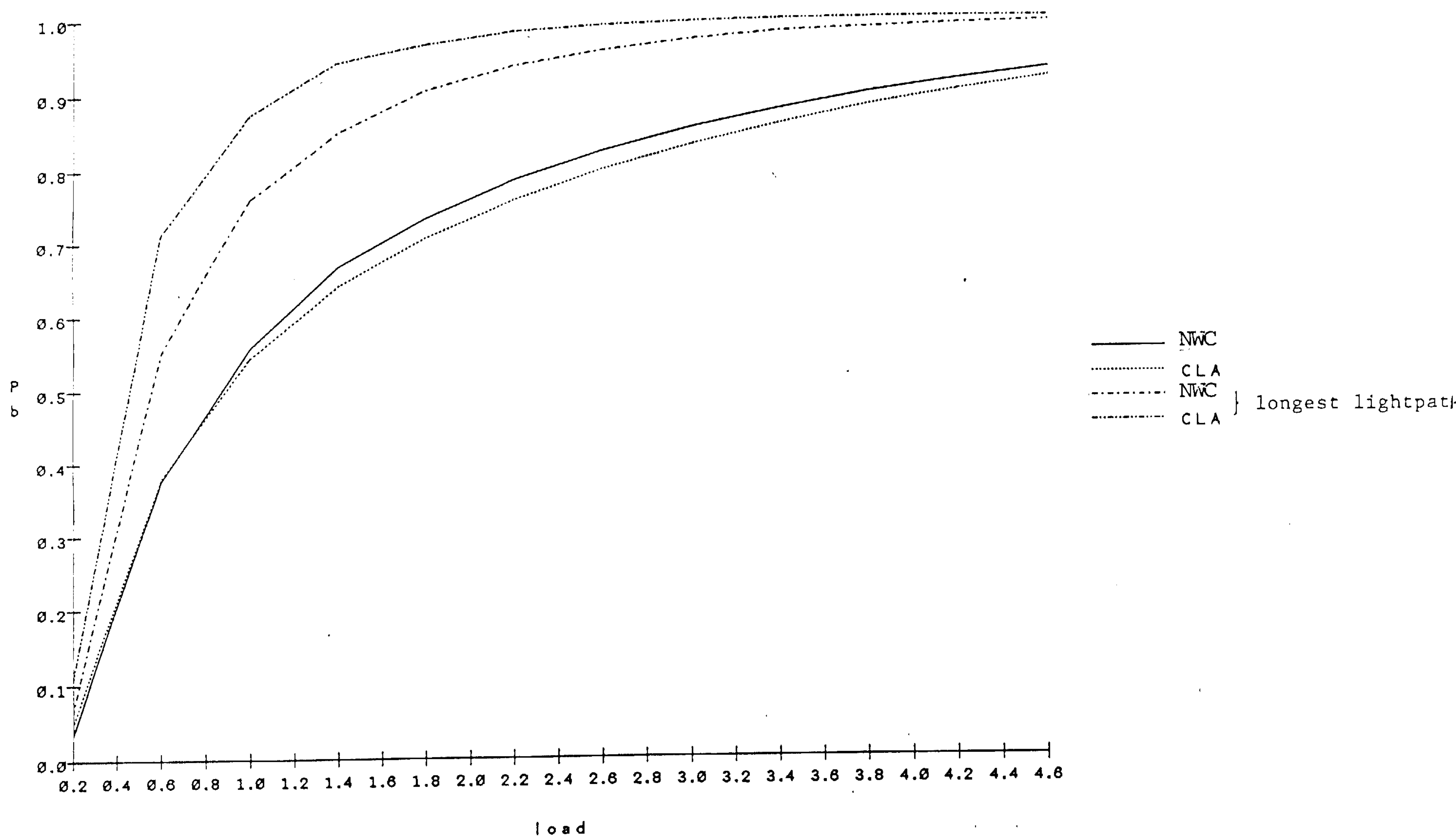


Figure 16 : Blocking Probabilities for CLA and NWC Algorithms for the Network in Figure 15

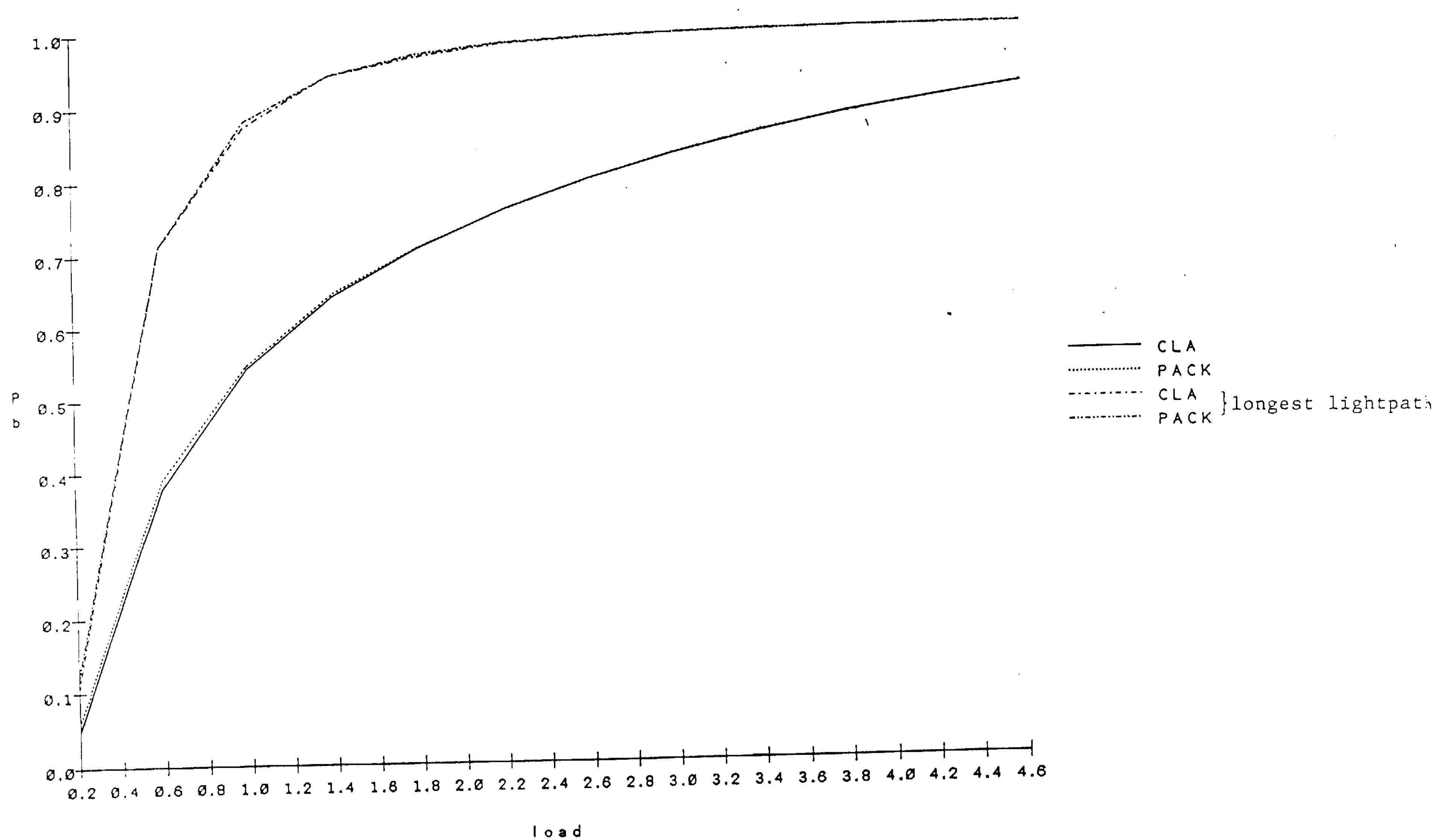


Figure 17 : Blocking Probabilities for CLA and PACK Algorithms for the Network in Figure 15

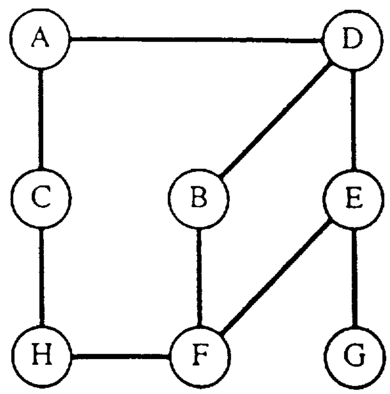


Figure 18(a) : Sample General Topology Network

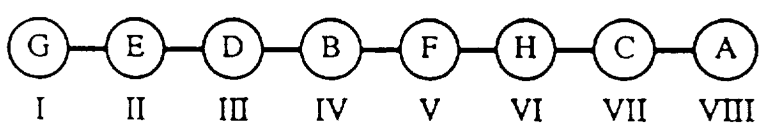


Figure 18(b) : String Representation

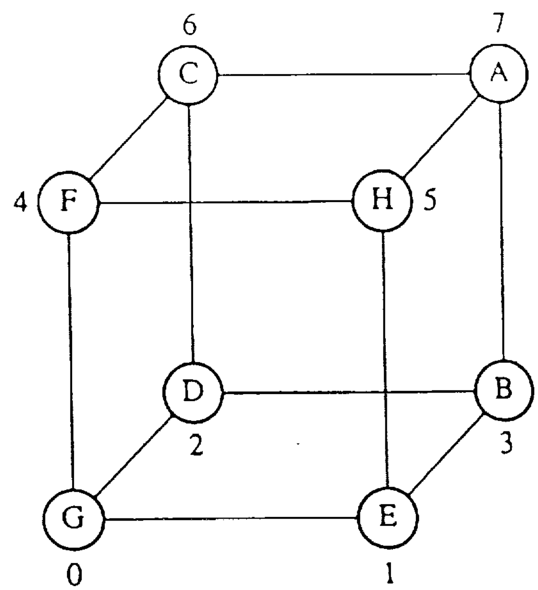


Figure 18(c) : An Eight Node Hypercube

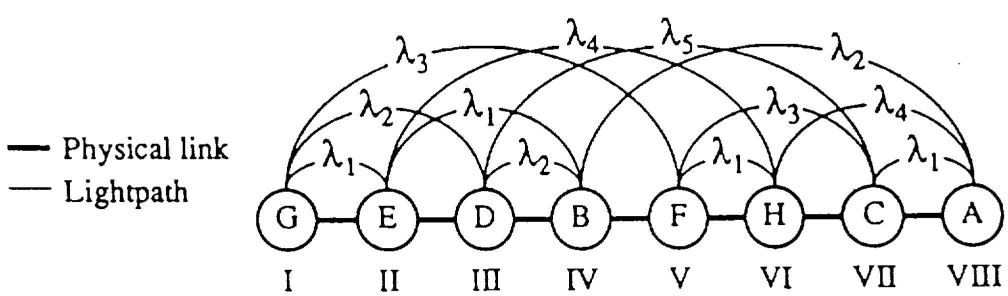


Figure 18(d) : Embedding of an 8 Node Hypercube in a String

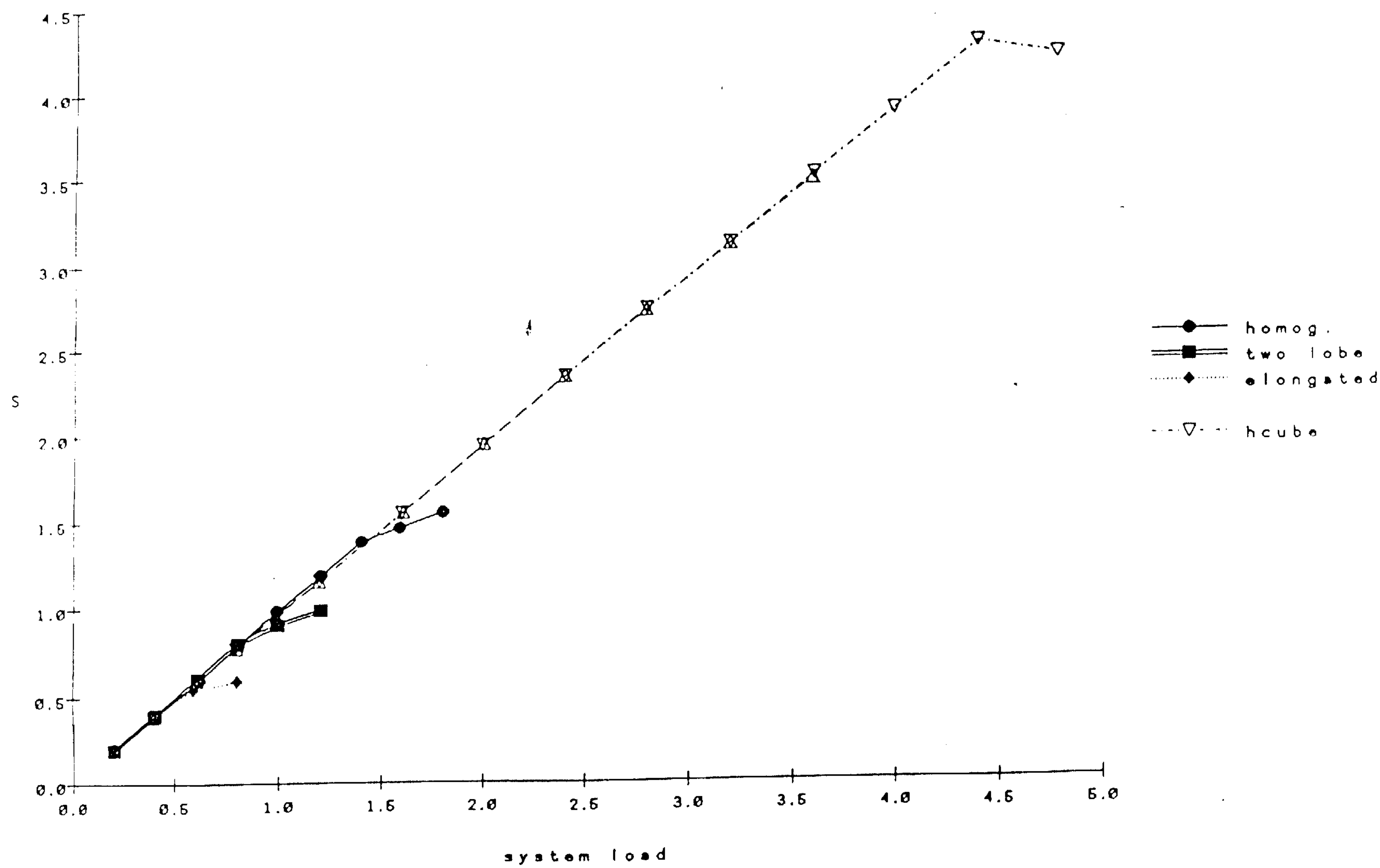


Figure 19 : Throughput vs. Load

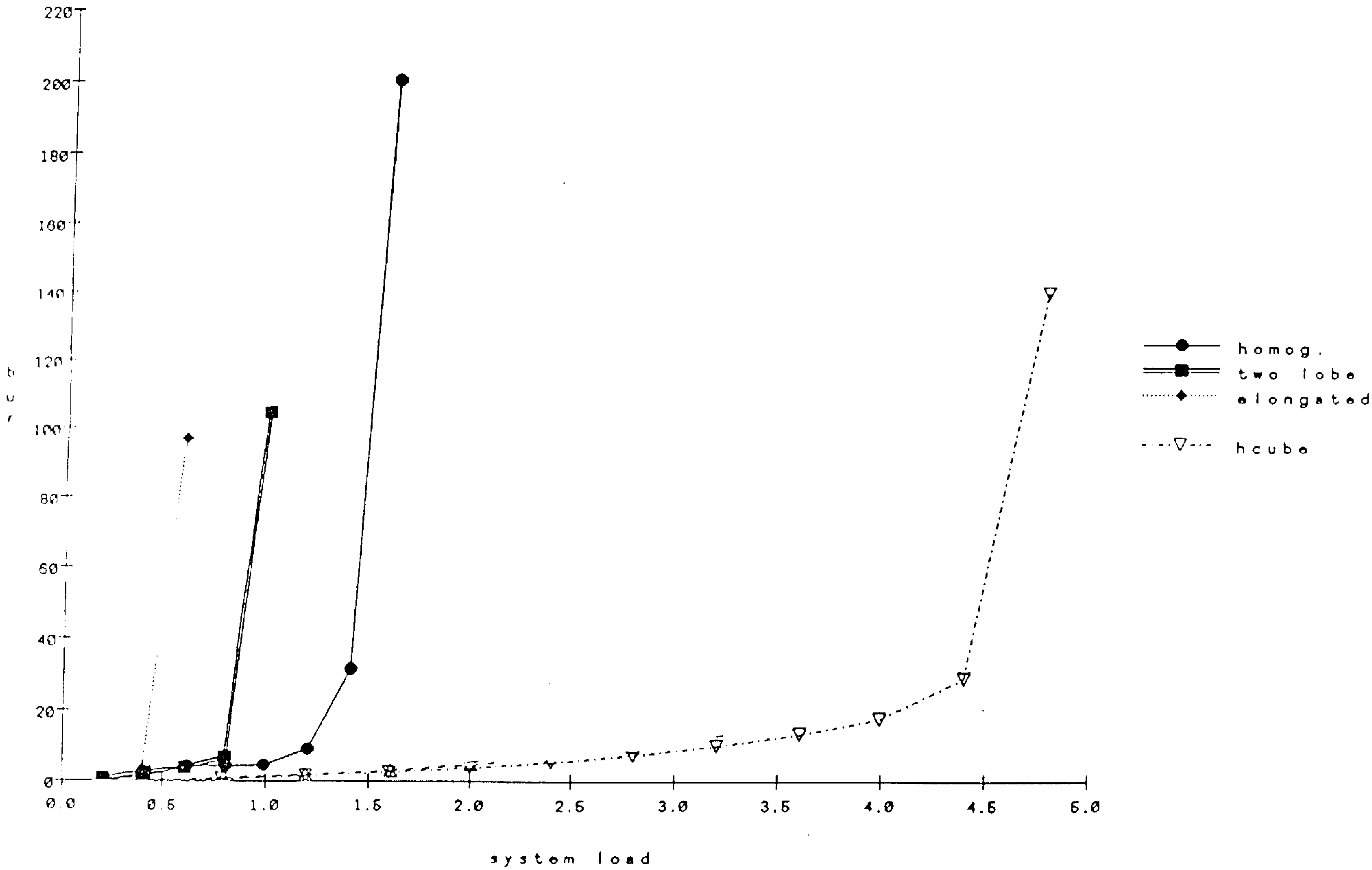


Figure 20 : Average Buffer Size (Maximum Loaded Node) vs. Load

topology	capacity	buffers		
		load=0.02	load=0.30	load=0.60
hcube	1.000	4.00	9.00	14.00
torus	0.641	5.00	11.00	51.00
De-Brujin	0.666	5.00	12.00	33.00
diamond	0.333	4.00	38.00	∞
htree	0.142	6.00	∞	∞
tree	0.058	8.00	∞	∞

Table 1 : Performance Measures for Lightnet with 128 nodes

	nodes			
topology	16	32	64	128
hcube	1.000	1.740	2.997	5.537
torus	0.997	1.537	2.479	3.535
De-Brujin	0.862	1.377	2.300	3.712
diamond	0.899	1.417	1.818	1.860
htree	0.585	0.603	0.731	0.772
tree	0.298	0.314	0.312	0.323
max-rnd	0.447	0.507	0.641	1.055
med-rnd	0.614	0.654	0.895	1.322
max-rnd4	0.112	0.127	0.160	0.264
med-rnd4	0.154	0.163	0.224	0.331

Table 2 : Capacity for various network sizes

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	5.10 ± 0.99	5.00 = 1.05	1.02
40	9.10 = 1.60	8.90 ± 1.52	1.02
60	11.50 = 2.07	11.30 ± 1.89	1.02
80	14.90 ± 2.33	14.80 ± 2.35	1.01
100	18.10 ± 2.60	18.10 = 2.60	1.00
120	22.10 ± 3.70	22.10 ± 3.70	1.00

(a) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 15

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	5.00 ± 1.25	4.90 ± 1.20	1.02
40	7.20 ± 1.14	7.10 ± 1.20	1.01
60	10.60 ± 2.01	10.20 ± 1.87	1.04
80	13.70 ± 1.70	13.60 ± 1.65	1.01
100	15.90 ± 2.73	15.50 ± 2.76	1.03
120	17.80 ± 1.62	17.40 ± 1.90	1.02

(b) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 30

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	4.60 ± 1.07	4.50 ± 1.18	1.02
40	6.90 ± 1.10	6.80 ± 1.03	1.01
60	9.50 ± 1.58	9.40 ± 1.65	1.01
80	13.00 ± 1.49	12.90 ± 1.79	1.01
100	15.50 ± 2.27	15.30 ± 2.54	1.01
120	16.20 ± 1.62	16.20 ± 1.62	1.00

(c) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 45

Table 3 : Effects of Network Size on the Number of Wavelengths
Required for the Static Unbounded Case

arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	6.92 ± 0.49	5.51 ± 0.44	1.26
0.40	11.90 ± 0.90	9.59 ± 0.85	1.24
0.60	16.49 ± 0.85	13.28 ± 0.54	1.24
0.80	19.90 ± 0.88	16.26 ± 0.89	1.22
1.00	25.90 ± 1.35	21.53 ± 1.33	1.20
1.20	28.95 ± 1.51	24.33 ± 1.40	1.19

(a) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 15

arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	6.39 ± 0.50	4.83 ± 0.48	1.32
0.40	10.58 ± 0.70	8.24 ± 0.66	1.28
0.60	14.06 ± 0.69	11.01 ± 0.72	1.28
0.80	17.87 ± 1.21	14.22 ± 1.17	1.26
1.00	21.51 ± 0.78	17.27 ± 0.86	1.25
1.20	24.75 ± 0.89	20.07 ± 0.89	1.23

(b) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 30

arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	5.88 ± 0.41	4.50 ± 0.41	1.31
0.40	9.68 ± 0.60	7.33 ± 0.68	1.32
0.60	13.09 ± 0.84	10.08 ± 0.66	1.30
0.80	16.89 ± 0.59	13.42 ± 0.60	1.26
1.00	20.64 ± 0.99	16.66 ± 1.03	1.24
1.20	24.29 ± 1.11	19.77 ± 1.16	1.23

(c) : Number of Wavelengths vs. Lightpath Request Rate for Network Size = 45

Table 4: Effects of Network Size on the Number of Wavelengths Required for the Dynamic Unbounded Case

APPENDIX A

submitted to

IEEE Transactions on Communication

Lightpath Communications : An Approach to High Bandwidth Optical WANs

I. Chlamtac   A. Ganz   G. Karmi
University of Massachusetts
Amherst MA 01003

Abstract

Emerging applications require a substantially higher bandwidth than the one offered by current networks. The technology necessary for providing the high bandwidth on the optical fibers, by means of Wavelength Division Multiplexing (WDM), exists. However, none of the network architectures proposed so far can efficiently tap this bandwidth in the wide area domain, due to the limitations imposed by the processing, buffering and switching required in these solutions. In this paper we propose a novel architectural approach that meets the high bandwidth requirements by introducing a communication architecture based on *lightpaths*, optical transmission paths in the network. Since lightpaths form the building block of the proposed architecture, its performance hinges on their efficient establishment and management. We show that although the problem of optimally establishing lightpaths is NP-Complete, simple heuristics provide near optimal solutions for several of the basic problems motivated by the Lightpath architecture.

1. Introduction

Current network architectures fail to meet the emerging integrated demands of communication applications. First and foremost, a substantial increase in network bandwidth must be provided to support applications such as HDTV, super–computer communications and video–conferencing [1,13]. Co–existing with these vast bandwidth consumers, there will continue to be applications with substantially smaller requirements. Thus, in addition to the need for high bandwidth, a bandwidth dynamic range of up to seven orders of magnitude must be contended with efficiently [1]. Reliability and availability will also become critical issues in future high speed networks carrying services previously supported by different networks. Clearly, the degree of reliability of the new network must be at least as high as that provided in the past by the network carrying the most stringent of the integrated applications. Finally, many of the emerging applications will present demands both for predictable service and on demand data delivery, leading to the requirement for integrating packet and circuit switched policies on the same network.

Currently, Wavelength Division Multiplexing (WDM) [2–4] offers a solution to the problem of providing the required bandwidth on optical links. However, the existing switching, processing and buffering technologies lag behind the transmission capabilities, turning the switching nodes into the foci of congestion. Therefore, the bandwidth provided by optical communication links cannot be readily translated into effective network throughput, i.e. user available bandwidth [5–13].

Packet switching solutions have traditionally been motivated by the need for efficient utilization of bandwidth at the expense of increased processing in the nodes. Today, leading approaches for wide bandwidth WANs are solutions based on packet switching, usually termed "fast packet switching" (also ATM, ATD) [14–18]. In these solutions packets are not required to wait and be error checked at intermediate nodes. However, buffering, E/O conversion of the packet header and routing oriented processing are required. Therefore, the node bottlenecks created by the discrepancy between transmission and processing/buffering capabilities are not removed, leading to networks with limited effective throughput.

In this paper we propose an innovative solution to the problem of supplying wide bandwidth to the users. We employ WDM not only to increase the user available throughput, but also to simplify switching. The use of WDM for switching purposes, is strongly motivated by considering current time division multiplexing standards for high speed WAN communication [19,20]. In these, the inherent correspondence between time slots and data channels is utilized so that no identification of the packet header is required for switching at intermediate nodes. This leads to practical and simple switching without the need for processing. Analogously, WDM possesses the inherent capability to identify data channels without processing, through the association of these with wavelengths, i.e. wavelength routing [2,4].

Consistently with these observations, the proposed architecture is based on the use of *lightpaths*. A lightpath is an all optical path (data channel) established between any two nodes in the network, created by the allocation of the same wavelength throughout the path. A lightpath requires no processing or buffering at intermediate nodes and potentially, no intermediate E/O conversions. Thus we propose to exploit the vast bandwidth attainable in multiple wavelength systems to establish a tradeoff between transmission bandwidth and user available throughput. By employing lightpaths to carry packets (and possibly, circuits), the total network processing and buffering requirements are reduced, when compared with a conventional store and forward network in which packets are processed at each intermediate node. This reduction is achieved at the expense of increased bandwidth consumption due to the fixed allocation of transmission paths (the lightpaths) as well as possibly transmitting packets on paths (the lightpaths) longer than the those dictated by a shortest path routing policy. By reducing processing requirements, the lightpath concept significantly alleviates the electronic bottlenecks allowing increased user available throughput. Employing lightpaths as the sole medium for all network communications thus presents significant advantages. However, practical limitations on the transmission technology and optical devices (transmitters, receivers, optical switches, etc.) restrict the number of available wavelengths, so that it is not possible to establish a lightpath between every pair of nodes. Therefore, only a selected set of nodes can be connected by lightpaths, leading to a new type of lightpath based networks, termed Lightnets. The Lightnet nodes correspond to the actual network nodes, while the links

correspond to lightpaths. The lightpaths thus serve as building blocks for the construction of Lightnet topologies, having as objective the minimization of the number of nodes actively involved in transmitting a packet, therefore minimizing the processing/buffering required to transmit a packet end–to–end. The topologies can be further optimized for routing, congestion control, or special reliability requirements.

Since lightpath are the basic building block of Lightnets, their correct and efficient establishment is crucial to the successful implementation of these architectures. Following an overview of the Lightnet architecture we, therefore, proceed to study the lightpath establishment problem in detail, analyzing its complexity and providing efficient solutions for it.

2. The Lightpath Architecture

We introduce the *lightpath* as an "optical communication path" between two (not necessarily adjacent) nodes, established by allocating the *same* wavelength throughout the route of the transmitted data. As a result, transmissions between lightpath endpoints require no processing or buffering at intermediate nodes. In this way, lightpath communication is targeted for implementation in an optical WDM network, addressing the mismatch between optical transmission rates and electronic processing speeds to alleviate the bottlenecks created at intermediate nodes.

To understand why the lightpath concept and the architecture built around it are inherently suited for high speed communication and to clarify the principles of Lightnet operation we consider the following analogy :

As a consequence of the need to transport more people and of evolving technologies, faster trains were developed, decreasing the amount of time passengers spent in actual travel. In this way, the bulk of total traveling time was shifted to the time spent at intermediate stops, increasing at the same time the congestion of intermediate train stations. The express train transportation system was developed to resolve these issues. When taking an express train, the passenger does not have to wait at intermediate stations and his travel time becomes determined primarily by the speed of the express train. Therefore,

the express system leads to lower passenger delays and alleviates the congestion problem, resulting in the need for smaller waiting room at intermediate stations.

The increased speed of optical communication (reduced packet transmission time) contrasted with the electronic processing rates at the switching nodes creates an apparent technological analogy: In the Lightnet solution, the wavelengths are the rails, a lightpath is an express train connection established between two stations and circuits/packets are the passengers. On the basis of lightpaths, the proposed architecture constructs an integrated packet and circuit switching solution : for packet switching, packets will be routed over "adjacent lightpaths", instead of being routed between all physically adjacent links, as in conventional store and forward packet switched networks. For circuit switching, bandwidth over multiple lightpaths can be allocated to a circuit for the circuit's duration.

To employ WDM to implement a network architecture based on lightpaths, a number of hardware related issues must be examined. Among them, the multiplexing technique and the associated number of channels made available, the photonic switches and the amplifiers required to overcome multiplexing, switching and path losses. With wavelength division multiplexing a limited number of high speed channels are available on each link. Current experimental systems are able to carry up to 20 channels, each modulated at 2 Gbit/s [21]. Similar devices approaching 60 channels are considered feasible in the near future [21,22]. Lightpath implementation at intermediate nodes requires the availability of suitable photonic switches. So far, emphasis on switching for LAN and WAN operation concentrated on switch operation with setup rates on par with packet transmission rates, a critical issue for packet switching networks. Electro–optic switches can be set up in less than 1ns [29], however, due to their poor crosstalk and attenuation characteristics, they may be impractical for an all–optical longhaul implementation. Mechanical optical switches of dimension up to 40x40, can switch single-mode signals with crosstalk of -90dB and attenuation of around 2dB, thus obtaining crosstalk and attenuation characteristics far better than those offered by electro–optic devices [23]. The use of mechanical switches, despite these characteristics, was not previously considered for data switching networks due to their slow set up speeds. The principle of using preset lightpaths changes this situation dramatically: while the setup time of the mechanical switches

is relatively large (50ms) [23], this does not constitute a problem in Lightnet, as lightpaths are not established on a per packet basis and can have lifetimes measured in hours or days. Furthermore, the design of the photonic switch required for lightpaths at the intermediate nodes can benefit from the wavelength continuity property of the lightpaths. We observe that since a lightpath maintains the same wavelength throughout its span, a channel incoming on one wavelength need not be switched to another wavelength. Consequently, in realizing the photonic switch, it is possible to *group* the channels according to wavelengths prior to switching. The photonic switch can thus consist of ω switching matrices, one for each wavelength. Each of these switches has dimension of $(D_p + D_n) \times (D_p + D_n)$, D_p being the physical node degree and D_n the number of lightpaths terminating at the node, as contrasted with a substantially more complex, $(\omega D_p + D_n) \times (\omega D_p + D_n)$ switch, that would be required without wavelength continuity. We also observe that the wavelength continuity of a lightpath ensures that no wavelength translation will be required within a lightpath. Lastly, an important aspect of the lightpath architecture is the end–to–end lightpath span. Experiments conducted recently have shown very successful use of optical amplifiers [24-29]. For instance, in [24], 25 optical amplifiers were used in series (an amplifier every 80km) to provide transmission of a 2.5 Gbit/s optical signal over 2,223 km of single mode fiber, with a power penalty due to accumulated noise of only 4.2dB. The combination of these hardware aspects of lightwave communication and the special properties of lightpaths suggest that a lightpath based Lightnet network can present a technologically feasible solution for a wide area wavelength routing network. The presented architecture thus carries a number of benefits :

- It can reduce the number of *active nodes* through which a packet is switched between source and destination (only the lightpaths' end nodes), thus alleviating the processing and buffering bottlenecks.

- The Lightnet presents a possible solution to congestion problems, fault conditions and provides a viable approach for networks having long term varying asymmetric traffic patterns. The capability to account for these issues in a Lightnet architecture is the result of the possibility to reassign lightpaths.

The switching nodes' hardware requirements are simplified, enabling the use of reduced

size *passive* devices due to the wavelength continuity of the lightpaths, as well as enabling use of photonic switches with slow set–up times. At the same time, notice that although the Lightnet approach allows for a linearization of the switching complexity as a function of the number of wavelengths, photonic switch sizes, as well as the current WDM state of art, impose a limitation on the number of wavelengths made available. Thus, the efficient assignment of wavelengths to lightpaths leading to a minimization of the number of wavelengths required is an important aspect of the Lightnet architecture. We turn to address this issue in the next section.

3. The Lightpath Establishment Problem

Two central issues must be addressed when studying the assignment of wavelengths to lightpaths. First, since wavelengths are a precious resource, it is necessary to establish lightpaths *efficiently* in terms of the total number of wavelengths required. Second, the requirement for establishing a lightpath using the same wavelength throughout its route, introduces a potential bandwidth loss when compared to a lightpath establishment in which the continuity constraint is not imposed. This loss can be perceived either as an increase in the number of wavelengths required to successfully establish a given set of lightpaths, or as an increase in lightpath blocking probability, if the number of wavelengths is limited. In providing *efficient* solutions for lightpath establishment, our objective will be to find algorithms that minimize this loss.

In deriving a lightpath establishment algorithm, we first analyze the complexity of an optimal assignment of lightpaths, introducing the following model. We represent the network by a triplet $G(V, E, W)$ in which V represents the set of N nodes, $N = |V|$, E represents the set of *directional* fiber links between nodes in V, (assuming $(u, v) \in E$ if and only if $(v, u) \in E \ \forall u, v \in V$) and W is the set of wavelengths on each link, $|W| = \omega$. It is assumed that ω is equal for all links. We define a *lightpath request* for connecting a given source / destination node pair by the links constituting a path between them. To establish a lightpath, it is necessary to find an unallocated, identical wavelength, on all the lightpath's links.

The problem we propose to study is the *correct* and *efficient* establishment of lightpaths. The correctness aspect of lightpath establishment must solve the problems of collisions : the simultaneous allocation of the same wavelength to more than one lightpath on any given link. In terms of efficiency, our goal is to maximize the utilization of wavelengths. Thus, we shall seek solutions that minimize the number of wavelengths used, for a given set of lightpaths or the lightpath blocking probability, for a given rate of lightpath establishment requests. We propose to achieve this goal by allocating wavelengths in such a way that, given the allocation of wavelengths to existing lightpaths, a maximum number of new lightpaths can be allocated. Figures 1*a* and 1*b* exemplify the lightpath allocation problem. The figures depict lightpaths establishment in a network with two available wavelengths ($\omega = 2$). In Figure 1*a* the allocation is done in such a way that any additional future lightpath establishment request can still be accommodated. In the allocation depicted by Figure 1*b*, if a lightpath request $v_1 \rightarrow v_3$ arrives before an existing lightpath is terminated, it will be blocked.

Definition : Static Lightpath Establishment (SLE) problem – given a network $G(V, E, W)$, $\omega \geq 3$, and a predefined set of lightpaths L, is it possible to establish all lightpaths in the set ?

We proceed to prove the NP-completeness of SLE by showing that the problem is equivalent to the n–graph–colorability problem [30,31]. That is, finding the minimal number of wavelengths that would accommodate the demands would amount to finding the chromatic number of some (general) graph, where the number of colors, n, corresponds to the number of wavelengths, ω.

Theorem 1 : SLE is NP-Complete.

Proof : See appendix A.

Thus, even if all lightpath demands were predetermined, we would have to search for a heuristic solution for all but trivial demand sets. In the next sections we therefore present a number of polynomial time heuristic solutions for basic lightpath establishment problems.

4. Heuristics for Lightpath Establishment

In this section we study heuristics solving the lightpath establishment problems motivated by the proposed architecture. The fundamental aspects of lightpath communications are covered by considering two possible objective functions : minimization of the required number of wavelengths and minimization of lightpath blocking probability. For these, the establishment can either be static, with all lightpaths predetermined, or dynamic, where lightpaths are established and terminated on–the–fly.

Before proceeding to describe solutions to the lightpath establishment problem we develop a lower bound on the number of wavelengths required by an optimal algorithm and show, that for a certain class of topologies, this lower bound is tight. Consider the number of wavelengths required to establish a given lightpath set without the wavelength continuity constraint. This number is given exactly by the number of lightpaths passing on the "busiest" link (i.e. the degree of edge congestion) and it is also, a lower bound on ω. We term this lower bound policy Non–Wavelength Continuous (NWC). For the following special case, this lower bound is tight :

Theorem 2 : For networks with acyclic topologies, if the NWC policy requires ω wavelengths to establish a given lightpath demand set, then there exists an allocation of wavelengths to lightpaths such that no more than ω wavelengths are required under the wavelength continuous lightpath policy (WC).

Proof : see appendix B.

We note that for networks with topologies containing cycles this bound is not tight as exemplified in Figure 2. Consider a 3 node ring with a lightpath demand set consisting of 3 lightpaths is shown. The lightpath set shown can be established using 2 wavelengths under NWC yet 3 wavelengths are necessary with the WC policy. Having developed a lower bound on an optimal solution, we now turn to presenting solutions for the different aspects of the lightpath establishment problem.

4.1 Static Demands, Unbounded Number of Wavelengths

The first case to be studied is the one corresponding to the Static Lightpath Establishment problem : a network in which a set of n lightpath requests is predetermined and the objective function is to establish all demands using a minimum number of wavelengths.

We use a greedy allocation heuristic which iteratively allocates a given wavelengths to all possible edge disjoint (i.e. non–colliding) lightpaths to whom a wavelength was not yet allocated. The procedure terminates upon allocating a wavelength to each lightpath. Using an intuition first observed in task scheduling problems, we first sort the lightpaths according to their respective lengths, and then try to allocate the wavelengths to the *longest* lightpaths first. Intuitively, a long lightpath is harder to establish, since an unallocated identical wavelength must be found on more links. Therefore, by establishing long lightpaths first, a better wavelengths re–use should be achievable, leading to an overall smaller requirement of wavelengths for a given lightpath set.

The exact description of the solution uses the following data structures :

lpcm[i,j]	: the lightpath collision matrix. *lpcm[i,j]* $=$ *1* if lightpaths i and j have a link in common (lightpaths collide)
lpnum[i]	: lightpaths, ordered by descending length
w	: wavelength number currently assigned
set[i]	: sets of lightpaths
s,e	: start, end pointers to current set
lambda[i]	: wavelength definition array. *lambda[i]* points to the first lightpath in *set* using wavelength i
lpok[i]	: flags indicating if lightpath i was already allocated
n	: number of lightpaths in set

or(set,s,e,lpnum[i],lpcm) : function; returns *true* if lightpath *lpnum[i]* has a link in common with the lightpaths in the set *set[s]..set[e]*, based on the lightpath collision matrix *lpcm*.

procedure static_establish

```
begin
  lambda[1] = w = s = e = 1
  for i = 1 to n do lpok[i] = false
  while (e < n) do begin (*)
    for i = 1 to n do begin
      if not lpok[i] then
      if not or(set,s,e,lpnum[i],lpcm) then begin
        set[e] = lpnum[i]
        e = e + 1
        lpok[i] = true
        e = e + 1
      end
    end
    w = w + 1
    lambda[w] = s = e;
  end
end
```

The following two lemmas prove the correctness of the algorithm :

Lemma 1 : The algorithm eventually stops.

Proof : Stems directly for the fact that whenever w is increased (thus creating an empty set of lightpaths), at least one lightpath can be established. Thus, for every time w is augmented, e is incremented at least once, eventually terminating the while loop.

Lemma 2 : No two lightpaths having a link in common are allocated the same wavelength.

Proof : Lightpaths having common links will not be assigned to the set corresponding to the same wavelength since the *or* function for them will return *true* and hence, s will be incremented between the assignment of the lightpaths to *set*, implying that the lightpaths will be established using different wavelengths.

The worst case is obtained when $lpcm[i,j] = 1 \;\; \forall\, i,j$ leading to an overall time complexity of $O(n^2)$.

4.2 Static Demands, Bounded Number of Wavelengths

Since the number of available wavelengths in WDM systems is expected to remain limited, it is of importance to study the problem of establishing a given set of lightpaths when the number of wavelengths, ω, is bounded by *omega*. We note that in this case, it is possible that lightpaths will be blocked. The objective function therefore changes in this case to minimizing the ratio of lightpaths rejected to lightpaths requested, defining for the static case the lightpath blocking probability. We note that an existing drawback of this objective function is that it does not differentiate between long and short lightpaths. Hence, a policy using this objective function, will in effect, discriminate against long lightpaths. The relative effects of blocking probability as a function of lightpath length are studied in section 5.

We note that the previous heuristic maximizes the use of every wavelength before proceeding to allocate a new one. Thus, in effect, it intuitively maximizes the number of unused wavelengths in the network in case their number is bounded. Noting that as long as there is an unused wavelength, the lightpath blocking probability will be zero, we employ a variation of this heuristic for the bounded wavelength problem.

As before, we shall allocate a given wavelength to all possible lightpaths that have not yet been allocated a wavelengths. However, the procedure will stop *either* if all lightpaths have been allocated a wavelength or the available wavelength pool has been exhausted. In addition, to allow for an unbiased study of the effect of blocking probability as a function of lightpath length, we avoid sorting lightpaths according to lightpath length as in the previous case. Thus, the heuristic remains unchanged, except the line marked by (*) in the algorithm which is modified to :

$$\textit{while } (e < n) \ \underline{\textit{and } (w < omega) \textit{ do begin}}$$

The correctness of the lightpath assignment still holds by virtue of lemma 2. For the termination we note that lemma 1 still holds, but in this case, since the number of wavelengths is bounded, the overall time complexity is reduced to $O(n \times \min(\omega, n))$.

4.3 Dynamic Demands, Unbounded Number of Wavelengths

The interest in studying the case in which lightpaths are established and terminated dynamically, stems from the fact that the Lightnet topology can be modified by reassigning lightpaths. By establishing the lightpaths dynamically, the Lightnet can be reconfigured for purposes of reliability, availability or even adaptation to long term traffic patterns.

We observe that in addition to the efficient use of wavelengths, the issue of *stability* becomes of primary importance in the dynamic case. Past experience with dynamic resource allocation suggests that lightpath allocation solutions might display a "fragmentation" problem in which, while wavelengths may be available on each link on a given path between a source and destination, the continuity constraint over the total path is not satisfiable. Hence it is important to establish whether a given allocation algorithm deteriorates over time as it does for example, in many memory allocation schemes.

We note, that the approach developed for the static cases maximizes the use of every wavelength it allocates before proceeding to allocate a new wavelength. We therefore pursue this approach for establishing lightpaths dynamically, as it intuitively leads to the maximal reuse of wavelengths, or in other words, should reduce fragmentation. The above approach can be mimicked in a dynamic environment by a greedy heuristic that establishes each lightpath using the first available wavelength. Thus, a new wavelength will be allocated if and only if a lightpath cannot be established using any of the wavelengths already in use.

We first consider the case of an unbounded number of wavelengths. In the exact representation for this solution we shall use the following data structures :

lightpath[id]: lightpath information record, holding the following fields :

- *path:* the links constituting the lightpath
- *len:* the lightpath length
- *wavelength:* the wavelength assigned to lightpath

busy[i,j] : *busy[i,j] = 1* if wavelength j is currently assigned to a lightpath passing through link i

path(s,d,vec,len) : procedure; returns a route for a lightpath from s to d

The links constituting the route are returned in *vec* and the route length in *len*

getid(id) : function; assigns a unique *id* to a lightpath

wave : index, used in searching for an available wavelength

The establishment procedure scans the matrix *busy* by columns (wavelengths) attempting to find a column where all the entries corresponding to the lightpath's links are zero (unused). If no such column is found among the wavelengths currently in use, the wavelengths counter, *wave*, is increased so as to allocate a new wavelength. Following are the procedures used to establish and terminate lightpaths :

```
establish(s,d,id)
(* establish a lightpath from s to d *)
begin
  path(s,d,vec,len)
  getid(id)
  lightpath[id].path = vec (* save path for hangup *)
  lightpath[id].len = len
  (* find wavelength in which to establish lightpath *)
  found = false
  wave = 1
  while not found do begin
    tmp = 0
    for i = 1 to len do tmp = tmp + busy[q[i],wave]
    if tmp = 0 then found = true
    else wave = wave + 1 (*)
  end
  lightpath[id].wavelength = wave
  (* update data structure - lightpath established on wavelength wave *)
  for i = 1 to len do busy[q[i],wave] = 1 end
```

Lightpath termination is taken care of by the following procedure :

```
terminate(id)
```

(terminate a lightpath *)*
begin
  for i = 1 to lightpath[id].len do
    busy[lightpath[id].path[i],lightpath[id].wave] = 0
end

The correctness of the solution is established by observing that before establishing a lightpath using a given wavelength, *establish* verifies that the wavelength is not in use on any of the lightpath's links. The process of establishing a lightpath terminates by either finding a wavelength that is in use in the network, but not on any of the lightpath's links, or by incrementing *wave* beyond the number of wavelengths currently in use, in which case *busy[i,wave] = 0* for any link i. Establishing a lightpath in a network in which ω wavelengths are currently in use is carried out in time $O(N \times \omega)$ in the worst case, where N is the number of nodes in the network.

4.4 Dynamic Demands, Bounded Number of Wavelengths

Last, we consider the problem of dynamically establishing lightpaths in a network in which the number of wavelengths is bounded by *omega.* As before, the problem is motivated by the limitation on the number of wavelengths.

Following the reasoning of the preceding unbounded case, we again employ a greedy approach. The wavelengths are tried in sequential order, establishing a lightpath by allocating it the first wavelengths that is not in use on any of the lightpath's links. However, in this case, as the number of wavelengths is bounded, lightpath requests may be blocked. The heuristic, therefore, proceeds as before except in this case, prior to increasing the number of wavelengths it is checked if the maximal limit has been reached. Thus, by changing the line marked by (*) in the previous heuristic to

else if wave ¡ omega then begin
wave = wave + 1

the solution for the bounded wavelengths case is obtained. The correctness of the solution

is established as in the case of subsection 4.3. The worst case time complexity is also given as before by $O(N \times \omega)$, where ω is the number of wavelengths in use, $\omega \leq omega$.

5. Results

In this section we study the performance of the lightpath establishment heuristics concentrating on the efficiency of wavelength allocation. Having proven that the exact solution is NP-Complete, comparison to exact results is not feasible for any networks of interest. However, as pointed out in section 4 by removing the wavelength continuity constraint from lightpath establishment a lower bound on the number of wavelengths needed is obtained. Thus, a comparison to the lower bound obtained by non–wavelength–continuous (NWC) case, can be made to evaluate the performance of the various heuristics as well as to determine the relative penalty imposed by the continuity constraint of the proposed WC lightpath establishment solutions.

The performance of the presented heuristics was derived by simulating general topology networks under varying traffic conditions and objective functions. All results were obtained with a confidence level of 95%. Lightpaths were randomly created choosing source / destination nodes according to a uniform distribution. The links constituting each lightpath were chosen following a shortest path policy, assuming all links to be of unit length, with random tie breaking rule. For the dynamic environments, lightpath arrival rate refers to the number of lightpath establishment requests per unit of time. An arrival rate λ is implemented in the simulation as an exponentially distributed lightpath request interarrival time with mean $\frac{1}{\lambda}$. Lightpath holding times were assumed to be deterministic, and equal to 200 time units.

In Table 1 we study the case of unbounded number of wavelengths by observing the average number of wavelengths required to establish a given lightpath set size (static demands) for three different network sizes. For each set size, the results presented are averaged over 10 different randomly generated lightpath sets. We observe that the results for the wavelength continuous lightpath establishment policy (WC) and the NWC lower bound are practically identical. This result can be explained by considering the implica-

tions of theorem 2 (section 4). The only discrepancy that may arise between NWC and WC can occur only when cycles are contained in the network graph, with the lightpath demand set also forming a cycle. However, the probability of such a structure occurring, given a lightpath set, is much smaller than the probability of multiple lightpaths passing through a link in the network. Hence, with high probability, the most congested link in the network determines the number of wavelengths required by the WC lightpath policy as well as determining the NWC lower bound. The study of the effect of network size on the number of wavelengths required to establish a given demand set, shown in Table 1, supports the above observation. In addition it shows, that as the network size increases, the number of wavelengths required for a given set size decreases. This is due to the fact that in a larger network there are fewer collisions between lightpaths for the same lightpath set size.

In Figure 3 a system with static demands and a topology depicted by Figure 4 where the number of wavelengths, ω is set to 5, is studied. The objective function in this case is the minimization of lightpath blocking probability given in Figure 4 as a function of the lightpath set size. The average blocking probability for WC, shown in Figure 3, is *lower* than the NWC lower bound, by up to 2%. This apparent contradiction is explained by observing that the WC policy exhibits a 4% higher blocking probability than NWC when considering long (i.e. equal to the network's diameter) lightpaths only (5 hops). Since a long lightpath takes up system resources that can be used by multiple short ones, an average lower average blocking probability results when long lightpaths are blocked. Figure 3 also shows the blocking probability for short (1 link) lightpaths confirming the above observation, noting that lightpath blocking probability for the WC policy is lower than that of the NWC case.

Before studying the actual number of wavelengths required to accommodate systems with dynamic demands and an unbounded number of wavelengths (ω), or the blocking probabilities in systems with dynamic demands and bounded ω, we first verify the *stability* of these results, as defined in section 4. Figures 5 and 6 study the stability of the heuristics for the unbounded and bounded wavelengths cases respectively. Figure 5 considers the stability of the proposed heuristics in an unbounded wavelength network by

plotting the number of wavelengths required to establish all demands for three different lightpath set sizes as a function of time. It is shown that following the transient phase, the average number of wavelengths does not increase over time. Similarly, Figure 6 shows the blocking probability as a function of time for the bounded wavelengths case ($\omega = 5$). We note that in this case the blocking probability also remains practically constant over time.

A dynamic establishment of lightpaths without having the ability to perform wavelength reallocation to already established lightpaths, can be expected to have a notable bearing on dynamic lightpath establishment heuristic. Table 2 displays the number of wavelengths required for the lightpath solution and the NWC lower bound as a function of lightpath request rates. Comparing the values corresponding to the lightpath establishment heuristic with the NWC allocation we observe that for high rates, less than 25% additional wavelengths are required on average to establish lightpaths for the same lightpaths request arrival rate. For low request rates, up to 32% more wavelengths are required for the lightpath case. The difference however, is more than offset by the fact that for these rates, a small (6–10) number of wavelengths are required. We also observe that the ratio between the number of wavelengths required under WC and the NWC case remains almost invariant for different network sizes. Last, when observing the same lightpath request arrival rate over different network sizes, the absolute number of wavelengths required decreases, for reasons similar to the ones stated in the static case.

Figure 7 studies the case were the number of wavelengths is set to 5. The objective function in this case is the minimization of lightpath blocking probability, depicted in these figures as a function of the lightpath request rate, comparing the presented heuristic with the NWC lower bound. We observe that the heuristic performs with a relatively small penalty relative to the optimum.

Finally, it is of interest to investigate the relative improvement obtainable by increasing the number of available wavelengths. Figure 8 depicts the lightpath blocking probability as a function of the number of available wavelengths, ω, for various lightpath request rates in a dynamic lightpath establishment environment. It is noted that for lower rates, a small increase of ω leads to a substantial reduction in blocking probabil-

ity, whereas high rates require a large increase in the number of wavelengths to obtain a similar blocking probability reduction. We further observe that the blocking probability for a request rate of 0.5 tends to zero for $\omega > 20$. With higher rates the blocking probability increases, reaching 0.47 for the same ω and a rate of 2.0. Thus, we conclude that small increases in the number of available wavelengths can prove substantial reduction in lightpaths blocking probability for small (less than 0.5) lightpath request rates.

6. Conclusions

In this paper we presented a novel network architecture motivated by recent developments in optical communications and targeted towards emerging wide bandwidth applications. Through the introduction of the lightpath concept, the proposed Lightnet architecture makes use of developing transmission and switching capabilities in the photonic domain to overcome the inherent limitations of electronics based networks by introducing the lightpath concept. Since the performance of this architecture is tightly linked to the efficient establishment of lightpaths, a detailed investigation of the lightpath establishment problem was conducted. The complexity of this problem was proven to be NP–Complete. Heuristics covering static and dynamic lightpath establishment, both for bounded and unbounded number of wavelengths were therefore presented and evaluated. It was demonstrated, that the requirement to use the same wavelength along the entire lightpath carries a limited performance penalty. The wavelength continuity in establishing lightpaths was shown, on the other hand to allow for a reduced switch size, and avoid the need for wavelength translation within a lightpath. The lightpath approach thus offers a high performance solution realizable with simpler and less expensive technologies.

Appendix A

Proof of Theorem 1 :

First we show that solving the n–graph–colorability problem would also solve SLE. Define an undirected graph $G_L(V_L, E_L)$ with a node $i \in V_L$ for every lightpath in L. Two vertices $i, j \in V_L$ have an interconnecting edge $e \in E_L$ if the respective lightpaths have at least one

link in common. A coloring of V_L with n or less colors, so that no two adjacent vertices have the same color, would yield a wavelength allocation in W where no two lightpaths having a link in common require the same wavelength. Thus, finding a feasible coloring would also yield a feasible wavelength allocation, answering SLE.

To complete the proof we show that solving SLE would also solve the n–graph–colorability problem, thus showing that finding a polynomial solution to SLE is unlikely. To show this, we describe a polynomial time algorithm that translates any graph into a network and an appropriate set of lightpath demands. Given a graph $G_C(V_C, E_C)$ we translate the coloring of G_C into a lightpath establishment problem as follows:

(1) create a node v_i^0 for every node $i \in V_C$.

(2) for every edge $e = i \rightarrow j \in E_C$:

create 4 new nodes x, y, v_i^k, v_j^l and directed edges

$v_i^{k-1} \rightarrow x, \quad v_j^{l-1} \rightarrow x, \quad x \rightarrow v_i^k, \quad x \rightarrow v_j^l$

Attach the mark i to edges going from/to v_i's, and $x \rightarrow y$.

Repeat similarly for the mark j.

The designation of a node in the new graph, v_i^j, stands for the $j'th$ replication of the node corresponding to node i in the original network, $j = 0..d(i)$ where $d(i)$ is the node degree of i. The construction is exemplified for a 4 node graph in Figures $9a$, $9b$. Figure $9a$ contains a graph for which the n–colorability problem is to be solved. Figure $9b$ illustrates its translation to a network, the number on the links being the marks. The lightpath demand set L is defined by the $|V_C|$ lightpaths where lightpath i requires use of all links having i as a mark. We note that the complexity of the algorithm is $O(|E_C|)$.

Lemma : A solution to the SLE with n wavelengths implies that the chromatic number of G_C is less or equal to n.

Proof : The lemma follows immediately from the construction. If the lightpaths can be established then there exists a function assigning a wavelength to each lightpath so that no lightpaths sharing a link are assigned the same wavelength. Since two lightpaths share a link if and only if the respective nodes in V_C are adjacent, this

implies the existence of a function assigning a color to each node in V_C, so that no two adjacent nodes are assigned the same color.

Appendix B

Proof of Theorem 2 :

Notation :

l_i – lightpaths. We denote arbitrarily the source node of a lightpath as its right node and the termination node as the left node.

e_i – the set of (unidirectional) links constituting lightpath l_i.

$c(l_i)$ – the collision set of l_i. $l_j \in c(l_i)$ if $e_i \cap e_j \neq \phi$.

Lemma : Let $\mathcal{L}$ be a set of lightpaths defined on a graph $G(V, E)$ where G is a tree. If $|c(l_i)| \geq n \ \ \forall l_i \in \mathcal{L}$ for some n then there exists an edge in E through which at least n lightpaths pass.

Proof : By contradiction, i.e. assume that every link carries less than n lightpaths. First we note that $|e_i| \geq 2 \ \ \forall l_i \in \mathcal{L}$ since if there exists a single link lightpath, that link will carry all the collisions of the lightpath, i.e. a total of n lightpaths. We define a *dividing link* x of a lightpath l_i as the rightmost link in e_i for which there exists a lightpath l_j, $l_j \in c(l_i)$ such that all links $e_i \cap e_j$ (all the common links of l_i and l_j) are to the right of x. We note that the dividing link of a lightpath always exists and it is never the rightmost link. Consider now some arbitrary lightpath l_1 and let x be it's dividing link. x induces a disjoint partition on $c(l_1)$ into two sets, R_1 and L_1 where R_1 contains lightpaths colliding with l_1 to the right of the dividing link only, and L_1 contains all other lightpaths in $c(l_1)$. We note that by the definition of the dividing link and from the assumption that no link carries n lightpaths, neither R_1 nor L_1 is empty.

Consider now the lightpath l_2, $l_2 \in R_1$ that was the lightpath that defined the dividing link of l_1. This lightpath also has a dividing link, partitioning $c(l_2)$ into L_2 and R_2. We note that

- $L_1 \not\subset L_2$ since $l_2 \in L_2$ but $l_2 \notin L_1$.
- $L_2 \not\subset L_1$ by the definition of the dividing link.

Finally, we observe that this partitioning process can proceed endlessly. For the $k'th$ partition, we have that the lightpath chosen for the partition, l_k, ensures that $L_{k-1} \not\subset L_k$. Furthermore, since the underlying graph is a tree, $l_k \notin L_i$ for any $i < k$. Thus we have an infinite sequence of sets, L_i whose union is infinite, contradicting the finiteness of the lightpath set $\mathcal{L}$.

Theorem 2 (from the paper) : *The number of wavelengths required to establish any set of lightpath demands under the lightpath policy is equal to the number required by NWC.*
Proof : By induction on the demand set size.
Base : For set size $S = 1$, $\omega = 1$ both for NWC and for the Lightpath Policy (WC). Similarly, for $S = 2$, if the lightpaths collide then there exists a link common to both and therefore, $\omega = 2$ for NWC. Clearly also, $\omega = 2$ for WC.
Inductive step : Assume $\omega_{\text{NWC}} = \omega_{\text{WC}}$ for all lightpath demand set of size k or smaller and consider a demand set of size $k+1$ requiring ω wavelengths under NWC. By lemma 1, there must exist a lightpath l_0, colliding with no more than $\omega - 1$ other lightpaths. Take l_0 out of the set. We now have a demand set of size k which can be established under NWC with ω wavelengths. Therefore, it can also be established under WC with ω wavelengths. Given a correct LP wavelength allocation we can now establish l_0 since it collides with no more than $\omega - 1$ other lightpaths and there are ω available wavelengths. $\blacksquare$

C.2

References

[1] E. Nussbaum, "Communication Network Needs and Technologies - A Place for Photonic Switching?" *IEEE Journal on Sel. Areas in Comm. joint with Journal on Lightwave Technology*, Vol. 6, No. 7, August 1988.

[2] G. R. Hill, "A Wavelength Routing Approach to Optical Communications Networks," *GLOBECOM 1988* Tokyo, Japan, 1988.

[3] H. Ishio, "Review and Status of WDM Technology and its Applications," *Journal on Lightwave Technology,* Vol LT-2, No. 4, August 1984.

[4] M. C. Brain and P. Cochrane, "Wavelength Routed Optical Networks Using Coherent Transmission," *ICC '88*, Philadelphia, June 1988.

[5] S. Personick, "Review and Fundamentals of Fiber Optics Systems," *IEEE Journal on Sel. Areas in Comm.*, Vol. SAC-1, No. 3, April 1983.

[6] P. Henry, "Introduction to Lightwave Transmission," *IEEE Communications Magazine*, Vol. 23, No. 5, May 1985.

[7] D. Keck, "Fundamentals of Optical Waveguide Fibers," *IEEE Communications Magazine*, Vol. 23, No. 5, May 1985.

[8] E. Basch and T. Brown, "Introduction to Coherent Optical Fiber Transmission," *IEEE Communications Magazine*, Vol. 23, No. 5, May 1985.

[9] J. S. Turner, "New Directions in Communications," *IEEE Communications Magazine*, Vol. 24, No. 10, October 1986.

[10] M. R. Finley, "Optical Fibers in Local Area Networks," *IEEE Communications Magazine*, Vol. 22, No. 8, August 1984.

[11] P. Kaiser, J. Midwinter and S. Shimada, "Status and Future Trends in Terrestrial Optical Fiber Systems in North America, Europe and Japan," *IEEE Communications Magazine*, Vol. 25, No. 10, October 1987.

[12] R.J. Boehm, Y.C. Ching, C.G. Griffith and F.A. Saal, "Standardized Fiber Optic Transmission Systems," *IEEE Journal on Sel. Areas in Comm.*, Vol. SAC-4, No. 9, December 1986.

[13] I. Chlamtac and W. R. Franta, "Rational, Directions and Issues Surrounding High Speed Computer Networks," *Proceedings of the IEEE*, January 1990.

[14] Z. Haas and D.R. Cheriton, "A Case For Packet Switching in High – Performance Wide Area Networks", *SIGCOMM '87* Vermont, August 1987.

[15] J. Turner, "Design of an Integrated Services Packet Network," *IEEE Journal on Sel. Areas in Comm.*, Vol. 4, No. 8, November, 1986.

[16] J. Coudreuse and M. Servel, "Prelude : An Asynchronous Time Division Switched Network", *ICC '87*, 1987.

[17] P. Newman, "A Fast Packet Switch for the Integrated Services Backbone Network," *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 6, December, 1988.

[18] H. Suzuki, T. Takeuchi, F. Akashi, and T. Yamaguchi, "Very High-Speed and High-Capacity Packet Switching for Broadband ISDN," *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 9, December, 1988.

[19] T1X1.4 contribution, "Rates and Formats for Optical Network Interfaces," T1X1.4/85-017.

[20] T1X1.4 contribution, "Synchronous Optical Network (SONET)," T1X1.4/85-005.

[21] S. S. Wagner, et. al., "Experimental Demonstration of a Passive Optical Subscriber Loop Architecture," *Electronics Letters,* Vol. 24, No. 6, March 1986.

[22] G.R. Hill, Personal Communication.

[23] DiCon Fiberoptics Inc., 950-C Gilman Street, Berkeley, CA 94710.

[24] R. Alferness, "Waveguide Electrooptic Switch Arrays," *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 7, August 1988.

[25] S. Saito, et. al., "An Over 2,200km Coherent Transmission Experiment at 2.5 Gbit/s Using Erbium-doped-fiber Amplifiers," *OFC'90 Postdeadline Papers, paper PD2,* San Francisco, January 1990.

[26] N. Edagawa, et. al., "12300ps/nm, 2.4 Gbit/s Optical Fiber Transmission Experiment using Er-doped Fiber Amplifiers and Penalty due to Transmitter Phase Noise," *OFC'90 Postdeadline Papers, paper PD3,* San Francisco, January 1990.

[27] H. Taga, et. al., "459km, 2.4 Gbit/s 4 Wavelength Multiplexing Optical Fiber Transmission Experiment using 6 Er-doped Fiber Amplifiers," *OFC'90 Postdeadline Papers, paper PD9,* San Francisco, January 1990.

[28] W. I. Way, et. al., "Distribution of 100 FM-TV Channels and Six 622 Mb/s Channels to 4096 Terminals Using High Density WDM and a Broadband In-Line Erbium-Doped Fiber Amplifier," *OFC'90 Postdeadline Papers, paper PD21*, San Francisco, January 1990.

[29] K. Kikushima, E. Yoneda, and K. Aoyama, "6-Stage Cascade Erbium Fiber Amplifiers for Analog AM and FM–FDM Video Distribution Systems," *OFC'90 Postdeadline Papers, paper PD22,* San Francisco, January 1990.

[30] Garey, D.S. Johnson "Computers and Intractability : A Guide to the Theory of NP – Completeness," *W. H. Freeman & Company* 1979.

[31] N. Christofides, "Graph Theory – An Algorithmic Approach," *Academic Press* 1975.

[32] R. W. Conway, W. L. Maxwell, and L. W. Miller, "Theory of Scheduling," *Addison–Wesley,* 1967.

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	5.10 ± 0.99	5.00 ± 1.05	1.02
40	9.10 ± 1.60	8.90 ± 1.52	1.02
60	11.50 ± 2.07	11.30 ± 1.89	1.02
80	14.90 ± 2.33	14.80 ± 2.35	1.01
100	18.10 ± 2.60	18.10 ± 2.60	1.00
120	22.10 ± 3.70	22.10 ± 3.70	1.00

(a) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 15

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	5.00 ± 1.25	4.90 ± 1.20	1.02
40	7.20 ± 1.14	7.10 ± 1.20	1.01
60	10.60 ± 2.01	10.20 ± 1.87	1.04
80	13.70 ± 1.70	13.60 ± 1.65	1.01
100	15.90 ± 2.73	15.50 ± 2.76	1.03
120	17.80 ± 1.62	17.40 ± 1.90	1.02

(b) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 30

set size	policy		WAC / NWC ratio
	WAC	NWC	
20	4.60 ± 1.07	4.50 ± 1.18	1.02
40	6.90 ± 1.10	6.80 ± 1.03	1.01
60	9.50 ± 1.58	9.40 ± 1.65	1.01
80	13.00 ± 1.49	12.90 ± 1.79	1.01
100	15.50 ± 2.27	15.30 ± 2.54	1.01
120	16.20 ± 1.62	16.20 ± 1.62	1.00

(c) : Number of Wavelengths vs. Lightpath Set Size for Network Size = 45

Table 1 : Effects of Network Size on the Number of Wavelengths Required for the Static Unbounded Case

arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	6.92 ± 0.49	5.51 ± 0.44	1.26
0.40	11.90 ± 0.90	9.59 ± 0.85	1.24
0.60	16.49 ± 0.85	13.28 ± 0.54	1.24
0.80	19.90 ± 0.88	16.26 ± 0.89	1.22
1.00	25.90 ± 1.35	21.53 ± 1.33	1.20
1.20	28.95 ± 1.51	24.33 ± 1.40	1.19

(a) : **Number of Wavelengths vs. Lightpath Request Rate for Network Size = 15**

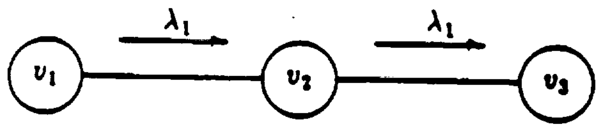
arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	6.39 ± 0.50	4.83 ± 0.48	1.32
0.40	10.58 ± 0.70	8.24 ± 0.66	1.28
0.60	14.06 ± 0.69	11.01 ± 0.72	1.28
0.80	17.87 ± 1.21	14.22 ± 1.17	1.26
1.00	21.51 ± 0.78	17.27 ± 0.86	1.25
1.20	24.75 ± 0.89	20.07 ± 0.89	1.23

(b) : **Number of Wavelengths vs. Lightpath Request Rate for Network Size = 30**

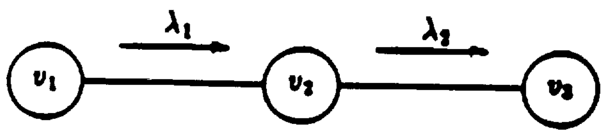
arrival rate	policy		WAC / NWC ratio
	WAC	NWC	
0.20	5.88 ± 0.41	4.50 ± 0.41	1.31
0.40	9.68 ± 0.60	7.33 ± 0.68	1.32
0.60	13.09 ± 0.84	10.08 ± 0.66	1.30
0.80	16.89 ± 0.59	13.42 ± 0.60	1.26
1.00	20.64 ± 0.99	16.66 ± 1.03	1.24
1.20	24.29 ± 1.11	19.77 ± 1.16	1.23

(c) : **Number of Wavelengths vs. Lightpath Request Rate for Network Size = 45**

Table 2 : Effects of Network Size on the Number of Wavelengths
Required for the Dynamic Unbounded Case



(a)



(b)

Figure 1 : Examples of Lightpath Allocation

$$L_1 : v_1 \rightarrow v_2$$
$$L_2 : v_2 \rightarrow v_3$$

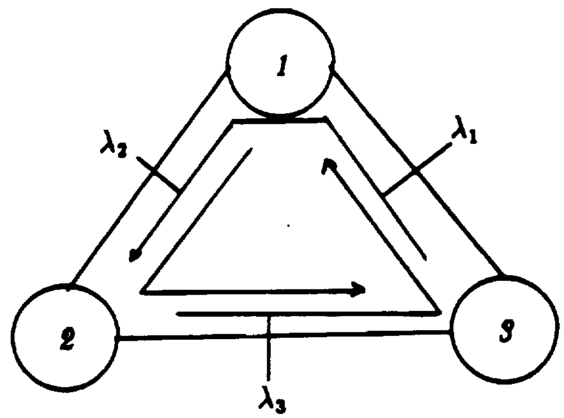


Figure 2 : Example For a Lightpath Demand Set in a Ring

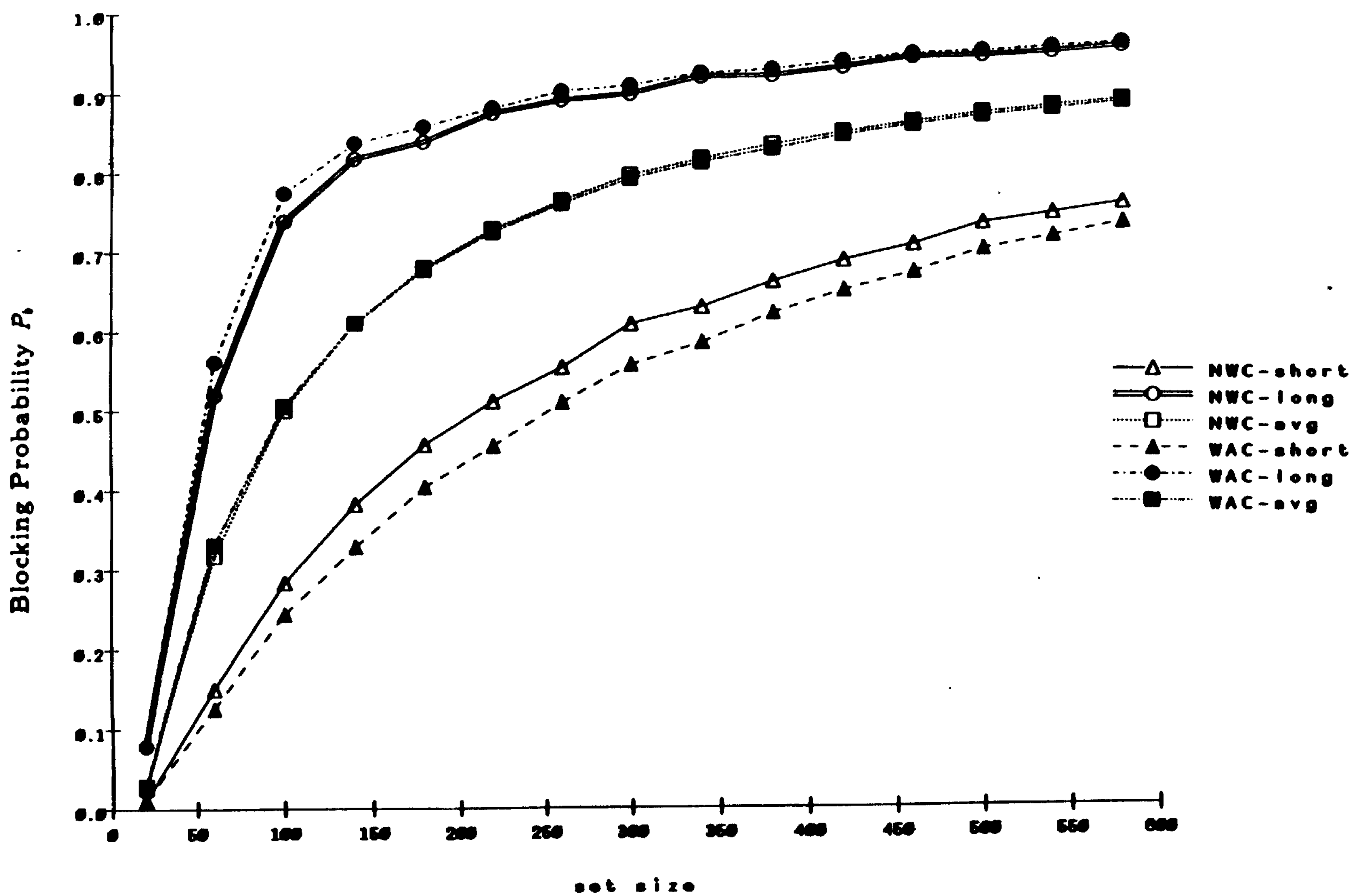


Figure 3 : Blocking Probability vs. Set Size (Static, $\omega = 5$)

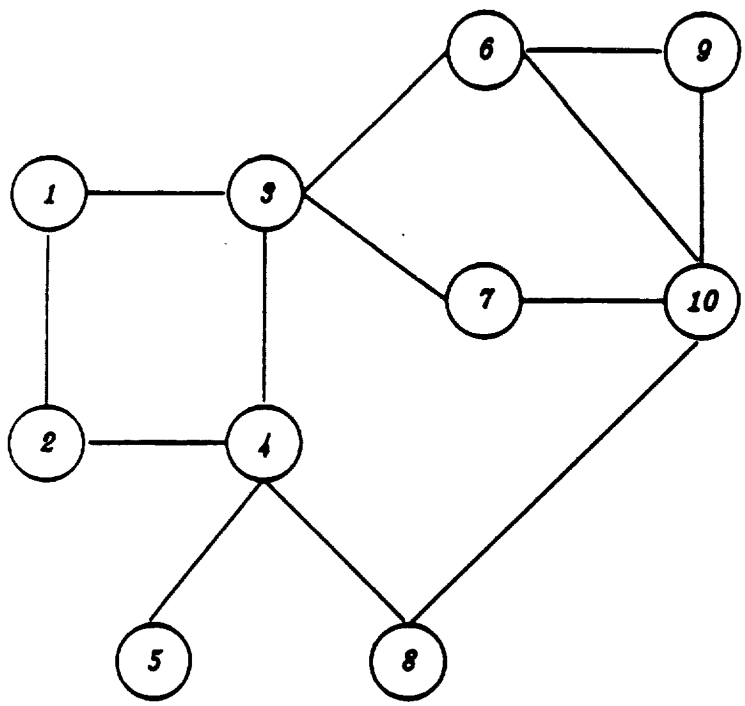


Figure 4 : Sample Network Topology

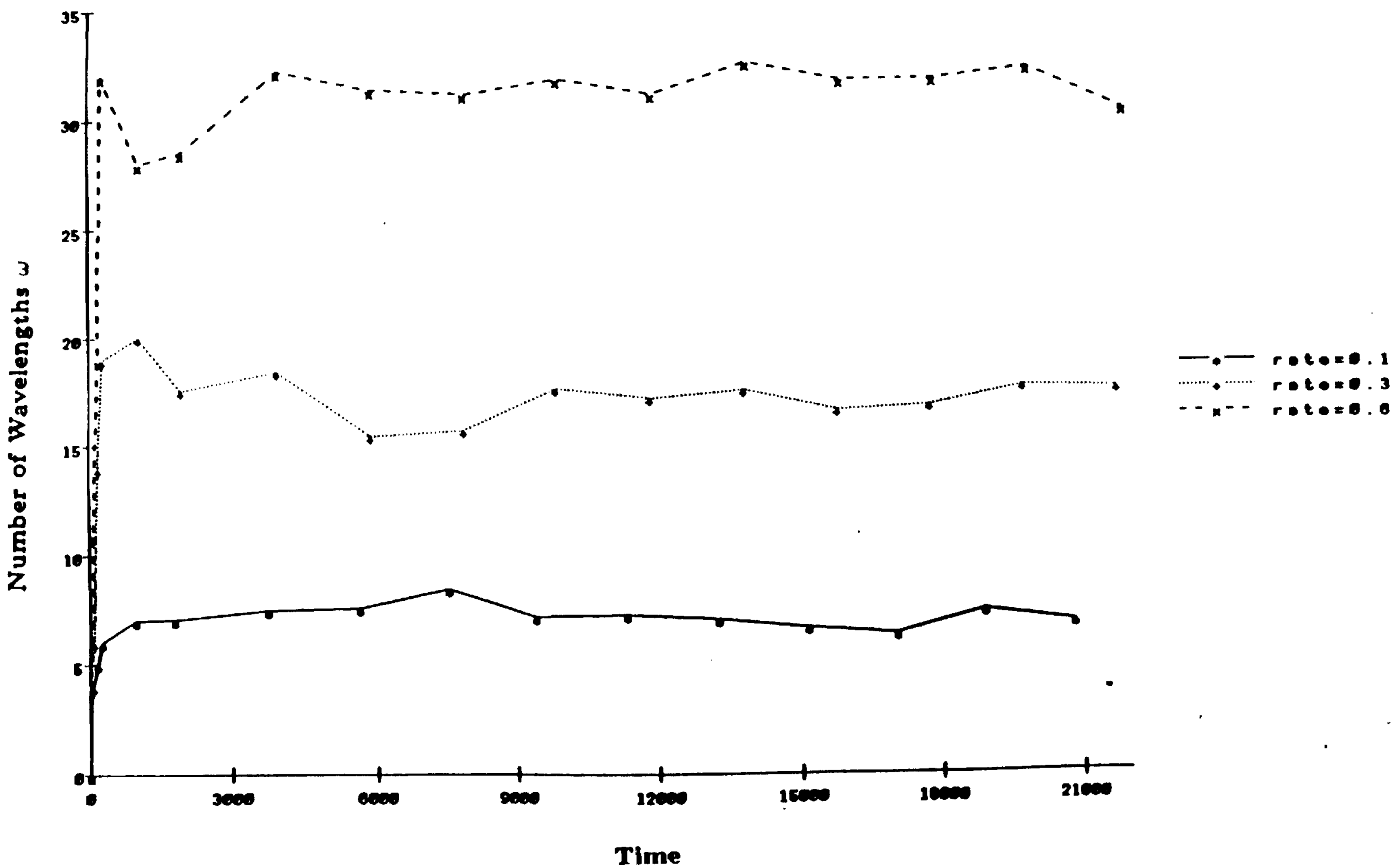


Figure 5 : Number of Wavelengths vs. Time (Dynamic, Unbounded)

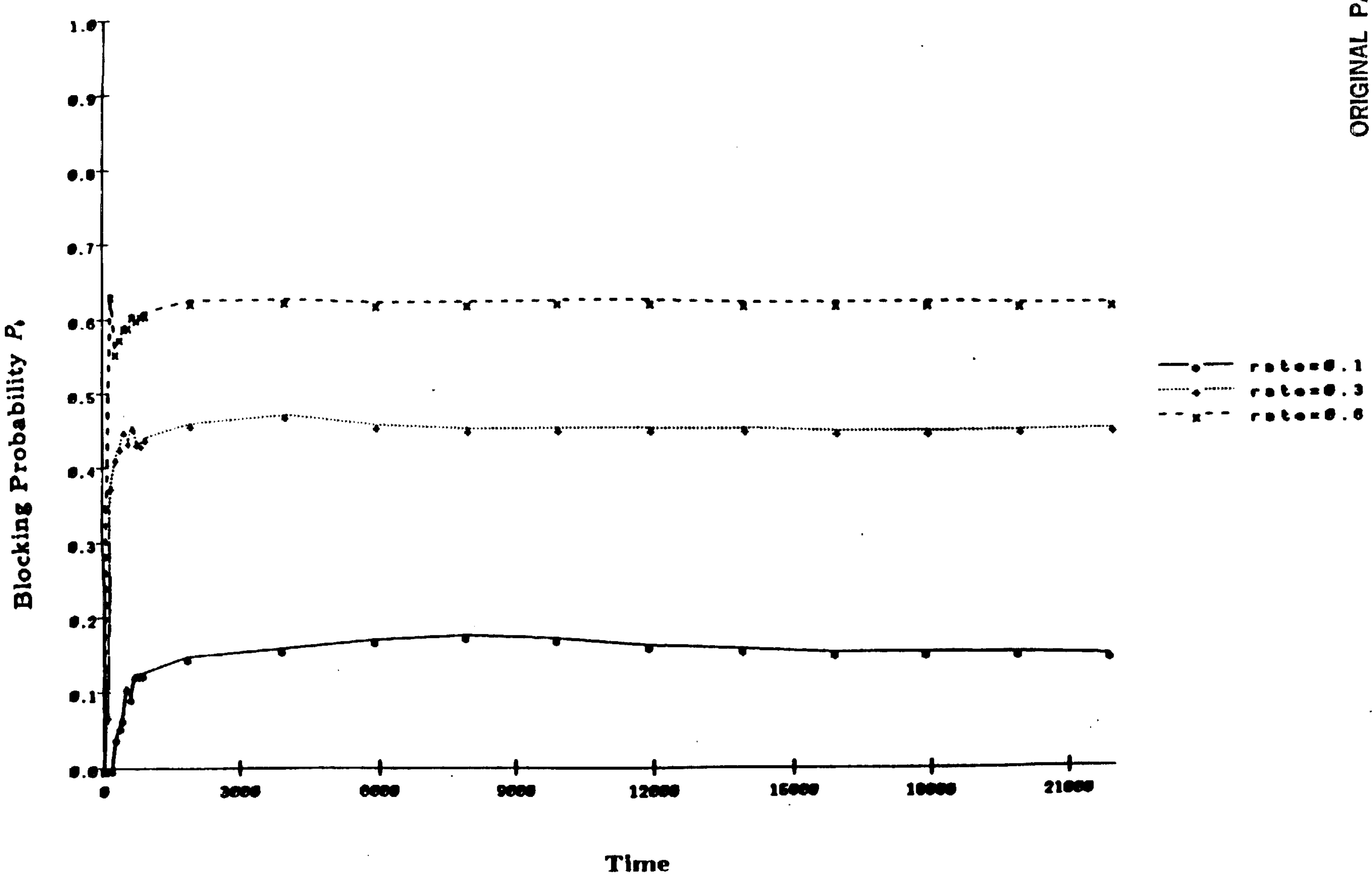


Figure 6 : Blocking Probability vs. Time (Dynamic, $\omega = 5$)

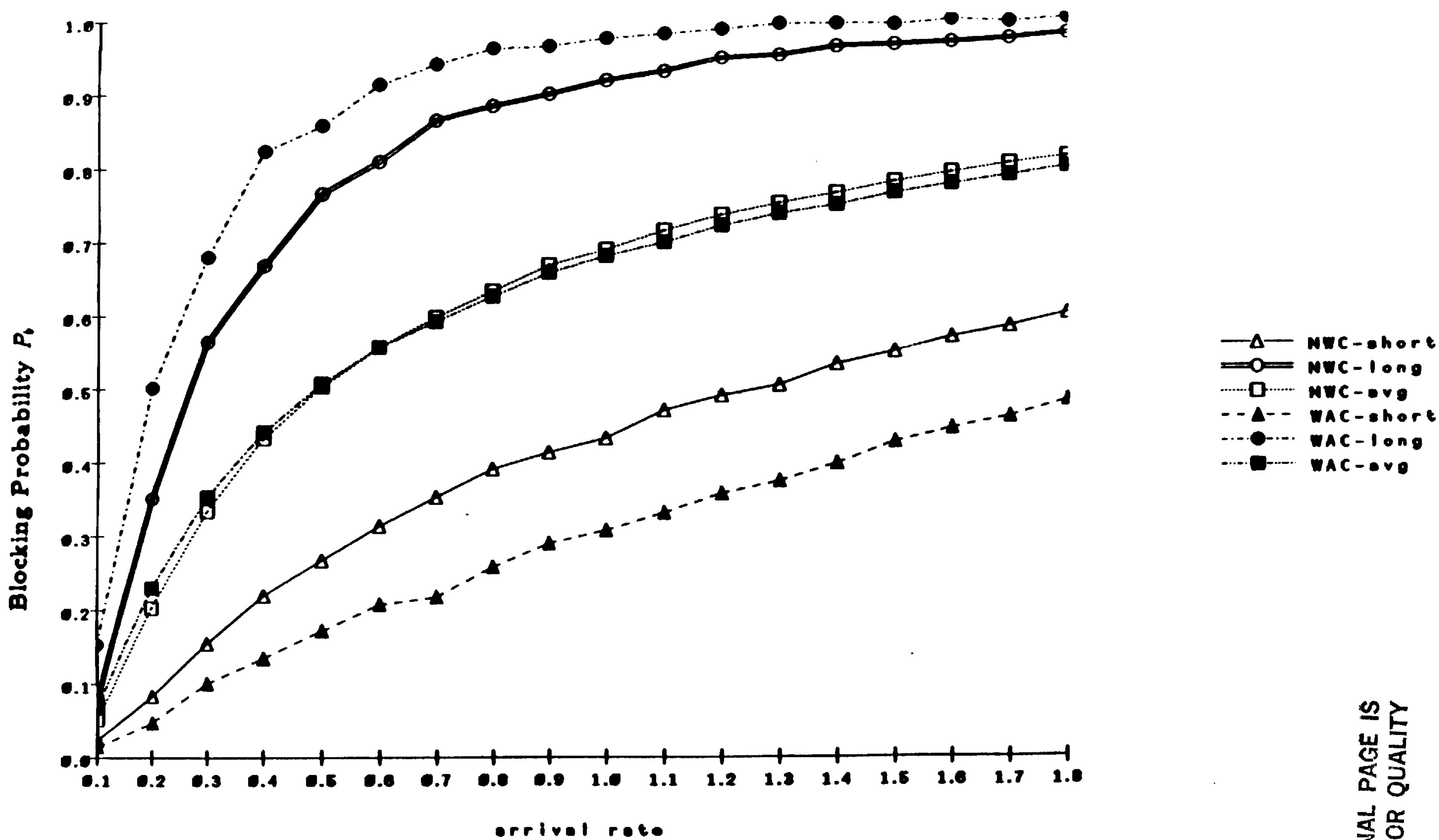


Figure 7 : Blocking Probability vs. Rate (Dynamic, $\omega = 5$)

ORIGINAL PAGE IS
OF POOR QUALITY

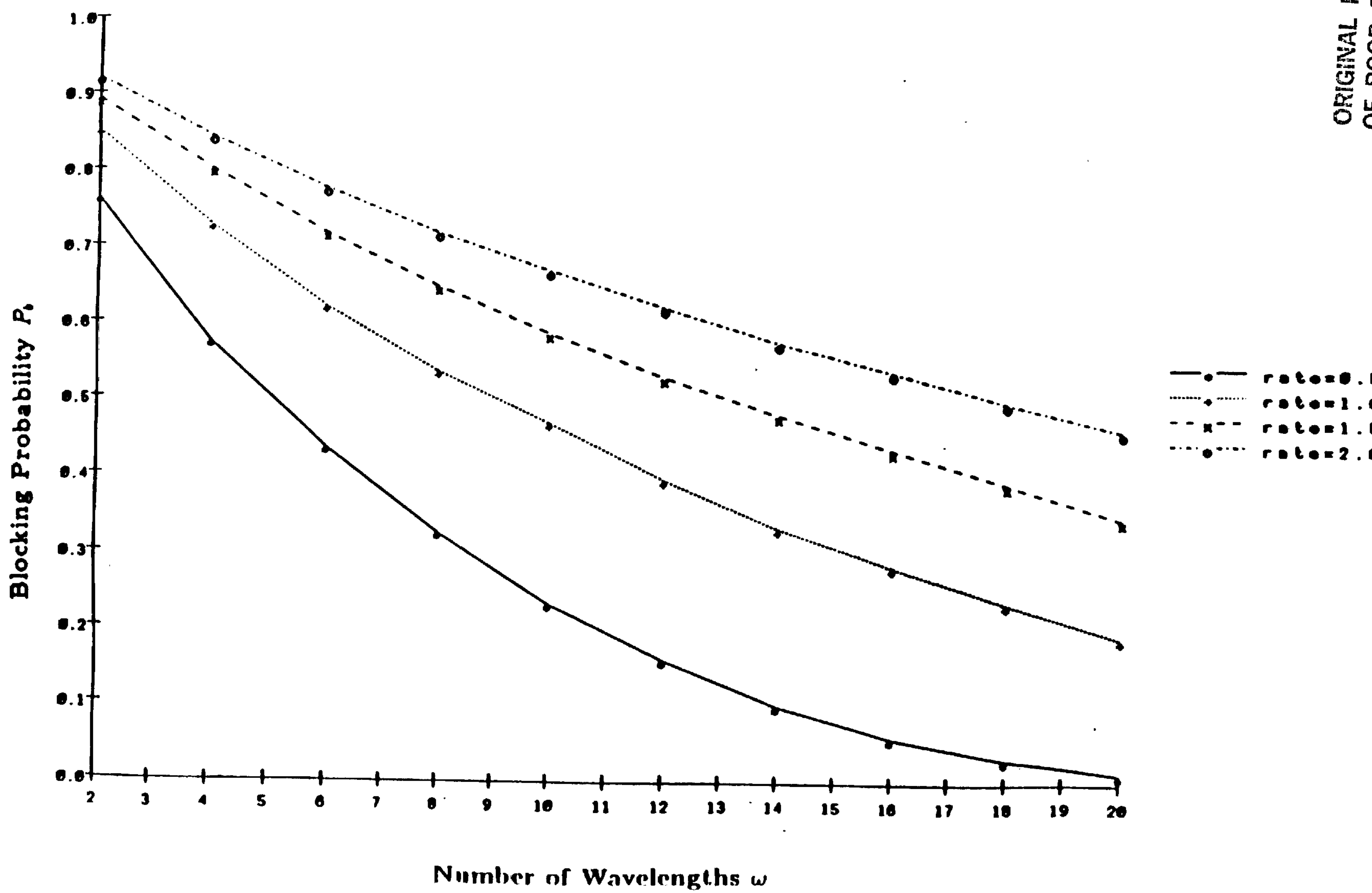


Figure 8 : Blocking Probability vs. Number of Wavelengths Bounded, Dynamic

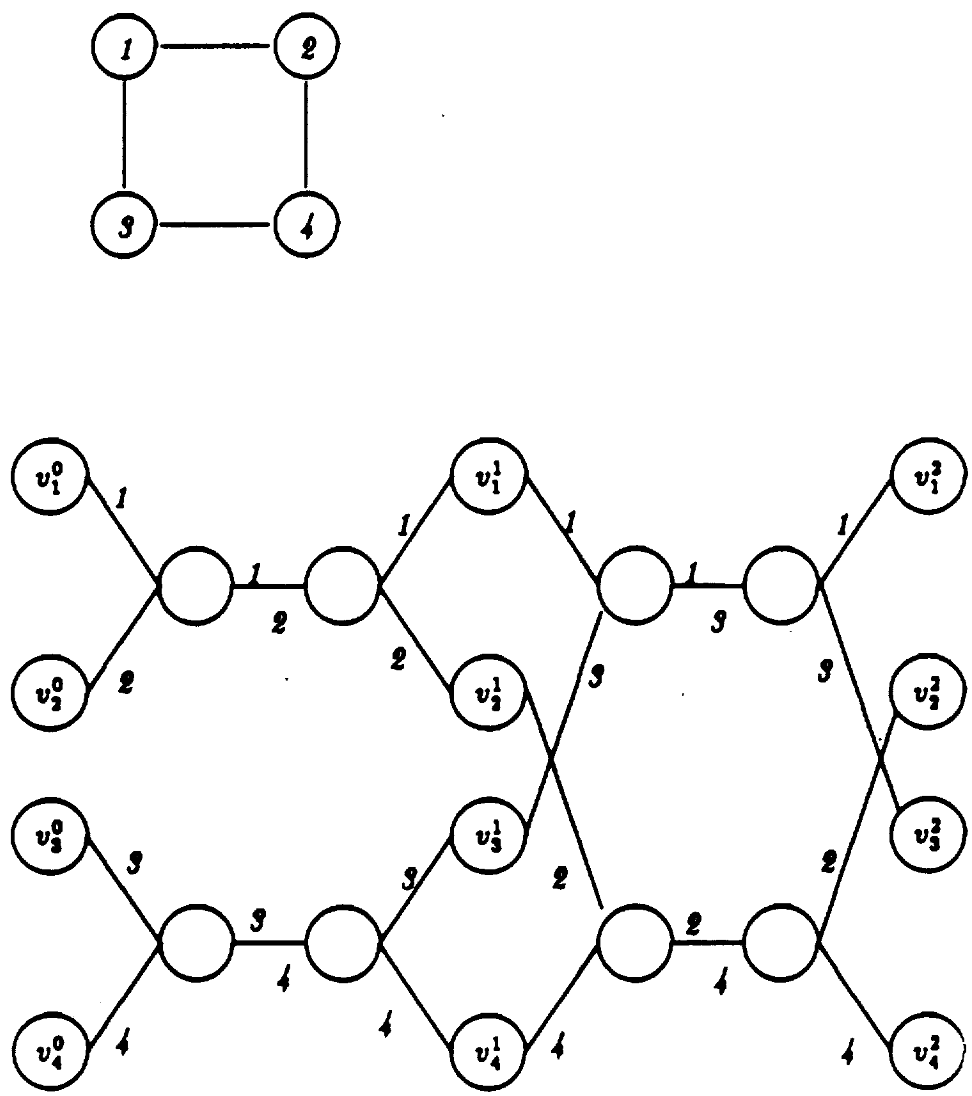


Figure 9 : (a) n–colorability graph

(b) Translation to SLE

APPENDIX B

presented at

Infocom '90
San Francisco, CA

Lightnet: Lightpath Based Solutions for Wide Bandwidth WANs

I. Chlamtac A. Ganz G. Karmi
University of Massachusetts
Amherst MA 01003

Abstract

An inherent problem of conventional point to point WAN architectures is that they cannot translate optical transmission bandwidth into comparable user available throughput due to the limiting electronic processing speed of the switching nodes. This paper presents the first solution to WDM based WAN networks that addresses this limitation. The proposed Lightpath architecture trades the ample bandwidth obtained by using multiple wavelength for a reduction in the number of processing stages and a simplification of each switching stage, leading to substantially increased throughput.

The principle of the Lightpath architecture is the construction and use of a virtual topology network in the wavelength domain, embedded in the original network. This paper studies the embedding of virtual networks whose topologies are regular, using algorithms which provide bounds on the number of wavelengths, switch sizes, and average number of switching stages per packet transmission.

1. Introduction

Wavelength division multiplexing is the emerging technology to provide very high bandwidth. Currently, however, there are no WAN architectures that can utilize the entire bandwidth offered by the optical fibers, supply wide bandwidth to the user and overcome the switching, buffering, and processing bottlenecks in intermediate nodes caused by the relatively lower speed of electronics. The purpose of the Lightpath architecture is to provide high speed integrated packet and circuit switching services to end users, while specifically addressing the growing discrepancy between processing and transmission bandwidths [1–3]. The Lightpath architecture is the first approach with the potential to achieve the above mentioned objectives, employing the emerging WDM technology to trade the ample optical bandwidth for a reduction in electronic processing/switching complexity, thus providing a direction for a realistic high speed network design.

The Lightpath architecture employs pre-established lightpaths – pre-established optical paths connecting two nodes via preset photonic space switches, as its basic transport vehicle. However it is practically infeasible to establish a lightpath between every two nodes in the network to create a clique, due to the total number of wavelengths needed and the need to terminate and process such number of high speed channels at each node.

The Lightpath architecture constructs a *virtual topology*, a *Lightnet*, in which lightpaths are the new links. Transmissions between any two nodes in the Lightnet take place on the lightpaths, passing on the way photonic switches within the lightpath and electronic switches between lightpaths. Lightnet topologies do not require a lightpath connection between every two nodes. Furthermore, we show that Lightnets remove the throughput limitations encountered when applying conventional WAN architectures to high speed networks. By using virtual topologies based on lightpaths, Lightnets 1) reduce the number of nodes actively involved in the routing/processing of a data unit, thus reducing electronic processing, buffering, and switching bottlenecks, 2) allow the link bandwidth utilization to exceed the electronic capacity of the adjoining nodes, and 3) make feasible, by being virtual, the setup and maintenance of regular topologies in the wide area domain.

The exact choice of the virtual topologies is therefore clearly a key aspect in the proposed approach. By using regular topologies it is possible to introduce inherent load balancing into the sys-

tem, simplify routing, congestion. and other control procedures, as well as determine the hardware requirements and performance. For instance, when using a virtual tree topology, the node degree η can be determined arbitrarily (e.g. $\eta = 3$ for a binary tree), while the total number of wavelengths is bounded by $\frac{1}{2}(\eta - 1)\log_{\eta} n$ for a network with n nodes [4]. On the other hand, a tree topology creates inherent traffic bottlenecks at all subtree roots with a potentially heavy performance degradation for most traffic patterns. In this paper we therefore choose the embedding of a hypercube regular topology, which does not incur this performance liability. The hardware resources required to embed a hypercube in terms of number of wavelengths and switch sizes are derived. Finally, the performance of a sample network is compared, operating under a conventional store and forward protocol and operating using the hypercube embedding.

2. Solution Overview

In Lightnet, a virtual topology is embedded in a general topology by associating the nodes of the virtual topology with the nodes of the original, general topology and implementing the edges of the virtual topology by creating lightpaths. The embedding process consists of three issues: first, the mapping of nodes in the virtual topology to nodes in the physical topology, determining a list of source / destination nodes in the physical topology that must be connected by lightpaths. Second, the determination of the physical links constituting each of these lightpaths, henceforth referred to as *lightpath routing*. Third, the allocation of wavelengths to lightpaths, so that 1) the same wavelength is allocated to a lightpath throughout its span, henceforth termed *wavelength continuity*, and 2) no allocation conflict occurs, a conflict being defined as the allocation of the same wavelength to two lightpaths passing through the same link. We observe that for a general set of source / destination nodes and a given lightpath routing, the wavelength allocation problem was addressed in [12]. In the context of regular topology embedding, we wish to solve the resulting instance of the wavelength allocation problem in a way which provides bounds on the number of wavelengths required, while exploiting the characteristics of regular topologies.

A polynomial time algorithmic procedure embedding a regular topology graph in a general topology network, while minimizing the number of wavelengths and consequently the complexity of the associated hardware, is not known [5,6]. We therefore introduce a two phase heuristic solution : in the first step we obtain a representation of any general topology network in the *simplest regular* form, a string, and then proceed in the second step to embed the regular topologies in this string.

The two phase approach is illustrated through the embedding of a hypercube in the general topology network of figure 1(a). For this network we first find an equivalent representation as a string, possessing the property that any two paths that are edge disjoint on the string are also edge disjoint in the original network. This property ensures that a wavelength allocated to two edge disjoint lightpaths in the string can also be allocated to the corresponding lightpaths in the original topology, without causing an allocation conflict. Figure 1(b) shows such a string representation. We notice that in this case a one to one correspondence exists between the edges of the string and those of the general graph. However, in the general case, a string edge may correspond to any subset of adjacent general graph edges.

We next establish a sequential mapping of hypercube nodes to string nodes, in which node 0 in the hypercube is mapped to node I in the string, node 1 in the hypercube to node II in the string, etc. Following this mapping, we obtain the lightpaths implementing the hypercube edges as shown in figure 1(d). Finally, an allocation conflict free wavelength assignment, using five wavelengths is also shown in this figure, i.e., all lightpaths passing on every physical link are allocated different wavelengths. The hypercube is drawn in the standard form in figure 1(c), noting that every *edge* in figure 1(c) corresponds to a *lightpath* in figure 1(d).

3. Obtaining a String Representation

We model a physical network topology as a directed graph $G_p(V_p, E_p)$ where V_p is the set of nodes and E_p the set of edges. Each edge carries ω wavelengths, determined as given in section 2, by the target, regular topology. We assume that if $(u,v) \in E_p$, $u, v \in V_p$ then also, $(v,u) \in V_p$.

Thus, transmissions on the same wavelength can proceed independently in opposite directions. For $G_p(V_p, E_p)$ we seek an *equivalent string representation*, $G_s(V_s, E_s)$ such that :

S1 $V_s = V_p$

S2 Each edge $e \in E_s$ connecting nodes $u, v \in V_s$ corresponds to a subset $\hat{E} \subseteq E_p$ forming a path in G_p from u to v.

S3 Any two paths that are edge disjoint in G_s are also edge disjoint in G_p where the edges are replaced by the corresponding edge subsets.

Conditions **S1–S3** guarantee that any regular topology embedding on the string G_s will also apply to the physical network G_p, using the *same* wavelength allocation. Condition **S3** further guarantees, that two lightpaths allocated the same wavelength in G_s can also be allocated the same wavelength in G_p. Therefore, bounds computed on the number of wavelengths and the associated node/switch capabilities needed in G_s, will also hold in G_p.

In the following subsections we present several approaches, carrying an inherent tradeoff between the conditions the graph G_p must meet vis à vis the amount of network resources required (number of wavelengths and passive switch size required to embed, in the second step, a regular topology in the generated string).

3.1 A Hamiltonian Solution

The most attractive approach in terms of hardware is obtained by generating the string through the identification of a Hamiltonian path. We observe that the process of generating a string from a Hamiltonian path is immediate as is the proof of the preservation of conditions **S1 – S3**. In such a path each node has a degree of 2, so that the size of the photonic switches in each node is minimal, with a total of ω 2×2 switches required, one for each of the ω wavelengths in the network.

If the physical layout of the original topology can be controlled, a Hamiltonian path can be established easily. Alternatively, it is known that the problem of finding a Hamiltonian path in a given arbitrary graph is NP–Complete, requiring therefore the use of heuristic solutions [7]. The solution presented in [8] presents an average polynomial time algorithm which finds a Hamiltonian path or establishes that none exists. For graphs created randomly, with a fixed probability of an edge existing between any two nodes, the algorithm was shown to find a Hamiltonian path if such exists with an average time of $O(|V|^3)$.

3.2 An Eulerian Solution

When a Hamiltonian path cannot be found, consider finding an edge disjoint path through all the nodes in the network, an Euler path, yielding a string in which all edges are traversed exactly once. As proven in [10] the string construction conditions **S1 – S3** are met for the Euler path solution. Compared to a Hamiltonian path, the presence of an Euler path is easily characterized; when such exists it is easily found [9]. On the other hand, each node may be traversed up to D times, D being the node degree (since each edge may be used once). Consequently, for this solution, each node will require ω photonic switching matrices, with the bound on the size of each matrix given by $D \times D$, ω being the number of wavelengths in the network as before.

The algorithm for finding an Euler path is a simple path extension algorithm [9]. Given the original graph G_p, first, find an Euler path, e.g. for the sample graph of figure 3, given by:

$$E \to D \to F \to B \to D \to A \to C \to H \to F \to E \to G$$

Second, tag the first instance of every node on the path. Replace all untagged nodes and the edges connecting them, between every two adjacent tagged nodes, by a single edge. For the above example this procedure generates the following string:

$$E \to D \to F \to B \to A \to C \to H \to G$$

3.3 A Spanning Tree Solution

This last string generation approach does not impose any conditions on the network topology beyond connectivity. It is based on finding a spanning tree of G_p. The string is then generated by a traversal of this tree. Since in this traversal the number of times each node is visited is bounded only by the physical node degree, the size of the passive switching matrices is again given by $D \times D$. To determine the number of switching matrices in each node we next present the detailed solution.

Let T be a directed spanning tree of G_p obtained by running a DFS algorithm on G_p [9]. A sample path described by the DFS algorithm for the graph given in figure 3 is given by:

$$B \rightarrow F \rightarrow E \rightarrow G \rightarrow E \rightarrow D \rightarrow A \rightarrow C \rightarrow H$$

On this path tag the first instance of every node and perform the replacement of edges, as described in the second step of the Euler path solution generating, for the previous path, the following string:

$$B \rightarrow F \rightarrow E \rightarrow G \rightarrow D \rightarrow A \rightarrow C \rightarrow H$$

As proven in [10], this construction meets conditions **S1** – **S3**. Note that in the preceding solutions, the edges in the string correspond to disjoint sets of edges in G_p. When generating a string by traversing a spanning tree, the corresponding sets in G_p are disjoint only if the edges have the same direction on the string (e.g. in the above example, the path from v_3 to v_5 along the direction of the string and the path from v_3 to v_4 in the opposite direction: both paths use the same edge, $v_3 \rightarrow v_4$ in G_p).

As a result, if ω wavelengths are required to establish a given topology on a string generated by either of the two previous solutions, 2ω wavelengths will be required when the string is generated from a spanning tree, a set of ω for each direction. This result establishes the size of the passive switch required for the spanning tree approach as 2ω switching matrices of dimension $D \times D$ in each node.

4. Embedding A Hypercube

In this section we present the hypercube embedding algorithm and then investigate the number of wavelengths required, the virtual network diameter, and the resulting switch size.

Embedding : Let $G_h(V_h, E_h)$ denote a hypercube with a node $v \in V_h$ numbered by an index i, $i = 0..n = 2^k - 1$, k integer and $n = |V_h|$. Number the nodes in a string G_s from left to right by a single index i, $i = 0..n-1$. Define the identity embedding function by :

$$\mathcal{E}(i) = i, \qquad i = 0..n-1 \tag{1}$$

Wavelengths allocation : Scan string from left to right. Define an ordering on the lightpaths based on their left end-node. Lightpath l_i is said to be smaller than lightpath l_j if it's left end-node is to the left of lightpath l_j's left end-node. The algorithm keeps track of the usage of wavelengths per link (a wavelength is used in a link if it was allocated to a lightpath passing through the link) and allocates wavelengths to the lightpaths based on the above ordering. Formally :

```
procedure alloc()
(* allocate wavelengths to lightpaths embedding
a hypercube in a string *)
(* w = (2/3)n : number of wavelengths required *)
begin
   for i := 1 to w do
        for j := 0 to n-1 do used[i,j] := false
(* used[i, j] = true if wavelength i was
allocated to a lightpath passing link j *)
   for i := 0 to n-1 do
        for all lightpaths with origin i do begin
             find a wavelength λ for which
                  used[λ,k]=false ∀ i ≤ k < d
(* d - lightpath destination *)
             allocate λ to the lightpath and mark
                  used[λ,k]=true ∀ i ≤ k < d
        end
end
```

Figure 1(d) illustrates the embedding of a hypercube using as the original topology the graph of figure 1(a). By observing the edges encompassed by each lightpath, as determined by figure 1(d), it is seen that the allocation is indeed conflict free, i.e., lightpaths having links in common have been allocated distinct wavelengths. As can be seen, a total of 5 wavelengths is required.

The properties of the hypercube embedding are as follows : a) a node degree which is logarithmic in the number of nodes, b) an average number of hops bounded by $\log_2 n$, and c) a linear number of wavelengths, given by $\frac{2}{3}n$. The first two properties are inherent to the topology chosen. In the remainder we establish the third property and determine its optimality.

Theorem 1 *For large n, the maximal number of wavelengths required when embedding an n node hypercube in a string with the identity embedding function and the above wavelength allocation algorithm is given by $\frac{2}{3}n$.*

Proof : We first proceed to prove a lemma counting the number of lightpaths passing through a link in the string :

Lemma 1 *For large n, the maximal number lightpaths passing through a link in the string is given by $\frac{2}{3}n$.*

Proof : Denote an edge $e \in E_s$ connecting nodes i and $i+1$ by $i+1$. We note that the number of lightpaths passing through an edge i is given recursively by

$$S(i,2n) = \begin{cases} S(n-i,n)+i & 1 \leq i < n \\ n & i = n \\ S(2n-i,n)+n-i & n < i < 2n \\ 0 & 2n \geq i \end{cases} \tag{2}$$

where $S(i,n)$ denotes the number of lightpaths passing through edge i in a hypercube with n nodes. We seek to find the maximum of this function. From [10] we have that

$$\max_{i\in[1.n-1]} S(i,n) \;=\; \frac{2}{3}n \tag{3}$$

We note that this lemma provides a lower bound on the number of wavelengths that will be required to implement the embedding using this embedding function. We next present a lemma motivating the simple algorithm given above for the allocation of wavelengths to lightpaths so as to attain the lower bound presented in the previous lemma.

Lemma 2 *Let L be a set of lightpaths defined on a string $G_s(V_s,E_s)$ and let the maximal number of lightpaths passing through any given link be m. Then, there exists an assignment of wavelengths to lightpaths using no more than $\omega = m$ wavelengths.*

Proof : See [10].

The proof of the theorem follows immediately from these two lemmas since no more than $\omega = \lceil \frac{2}{3}n \rceil$ lightpaths pass through any given link in the string and by virtue of lemma 4 an allocation of wavelengths to lightpaths exists which avoids collisions using no more than ω wavelengths. ♠

The correctness of the wavelengths allocation algorithm, and specifically the availability of a wavelength for allocation at the inner loop are guaranteed by the proof of lemma 4. The near-optimality of this embedding is proved by developing a lower bound on the number of wavelengths required to embed a topology G_e on another given topology G_g. We use a proof along the lines presented in [11]. Define a $\frac{1}{2}$-partition of a graph G as a partition of the node set of G into two equinumerable sets. In this partition an edge is said to be *comprehended* by one subset if both its endpoints are in the subset. Otherwise we shall term the edge a *cross-edge*. The $\frac{1}{2}$*-width* of G, $w(G)$ is defined as the minimal number of cross–edges where the minimum is taken over all possible $\frac{1}{2}$–partitions of G. Using these definitions we have:

Lemma 3 *A lower bound for the number of wavelengths required to embed an arbitrary topology G_e on another topology G_g is given by :*

$$\omega \geq \lceil \frac{w(G_e)}{w(G_g)} \rceil \tag{4}$$

Proof : See [10].

Theorem 2 *The embedding presented for the hypercube is optimal up to a constant factor.*

Proof : We first proceed to find a lower bound on the $\frac{1}{2}$-width of the hypercube.

Lemma 4 *The $\frac{1}{2}$-width of the n–node hypercube is given by $w(G_h) = \frac{1}{2}n$.*

Proof : See [10].
The theorem now follows since it was shown that at least $\frac{1}{2}n$ wavelengths are required while an algorithm employing $\frac{2}{3}n$ wavelengths was shown. ♠

5. Performance

To establish the efficiency of the Lightnet architecture we proceed to compare the performance observed using conventional, store and forward wide area network operation with the performance of the same networks employing the hypercube Lightnet embedding. In this embedding the number of wavelengths required is determined by theorem 1. In the embedding process the lightpath routing was restricted to the string, i.e. only the string edges are used to construct the lightpaths. The performance

measures studied are network capacity, the maximum network throughput that can be sustained while maintaining ergodicity in all queues and average buffering requirements at the maximally loaded node.

In the network model we assume : Packet processing takes place at all nodes on a packet's path in a conventional network and nodes performing switching between lightpaths in the Lightnet. A minimum-hop shortest path routing (in terms of nodes performing switching) with random selection for tie breaking rule is used. Since packet processing in high speed networks is significant, the node capacity is modeled as finite. A packet arriving at a node may enter a processing server if one is available, or join a common queue if all servers are busy. Considering the large discrepancy between the optical transmission bandwidth and the processing and propagation times, the former is considered negligible in the model.

To evaluate the network performance a simulation was developed. In it, packet arrivals are assumed to follow a Poisson distribution and source and destination selection for each packet follows a uniform distribution. Node processing capability was modeled by 3 parallel (packet) servers, each with service rate of 0.1 packet/unit time. Upon terminating service the packet proceeds to the next node. A propagation delay of 100 time units is assumed between physically adjacent nodes. All simulation results were obtained with 97% confidence levels.

The hypercube topology was embedded in three randomly created physical topologies : The first "*homogeneous*" topology was generated with a diameter and node degree distribution matching those of an earlier 1978 Arpanet topology. Two other topologies were introduced to study the relative effects of various bottlenecks on the performance of the store and forward and Lightnet architectures. The "*two-lobe*" topology consists of two clusters of 31 nodes each, randomly connected using the same parameters as the homogeneous network. To this configuration two nodes were added, each with one link to one of the nodes in each cluster. The third, "*elongated*" topology, was generated with a longer diameter of 20 as compared to 15 for the two lobe topology and 10 for the homogeneous topology.

Since the primary issue raised in the design of optical networks is the discrepancy between transmission which in turn leads to bottlenecks and reduced user available throughput, we concentrate on comparing network capacities. Figure 2 presents network throughput as a function of system load for the three physical topologies. The results show that in the random topology network capacity was nearly tripled when using the hypercube embedding. For the two-lobe topology, the hypercube embedding provided a capacity increase by a factor of 4.5. Last, in the elongated topology, network capacity increased by a factor of 7.5 for the hypercube embedding. The superior performance of the Lightnet embedding is due to the reduced number of active switching stages per packet transmission and the inherent load balancing when compared to the conventional network operation. Notice that the throughput of Lightnet is independent of the physical topology carrying the embedding. Although, for the same regular topology, packets will traverse different physical paths in different underlying physical topologies, the data paths in terms of nodes performing switching between lightpaths, the factor affecting capacity, remain invariable. Figure 3 shows the buffering requirements, demonstrating that with the Lightnet approach the increased network throughput does not require additional buffers.

6. Discussion and Conclusion

The purpose of this paper was to demonstrate a new approach to harnessing the emerging WDM technology for high speed WAN communication. To be effective, an optical wide area network must deal with the mismatch between the electronic processing rates and the optical transmission bandwidth and must present a solution with emphasis on viable optical/electronic switching. The principle of the introduced approach is the trading of the high optical bandwidth for reduced electronic processing/switching requirements, while providing for simplified optical switching at intermediate nodes. Utilizing the ample bandwidth provided by WDM in the form of multiple parallel channels, the Lightpath architecture introduced a number of new ideas which converge into such solution: 1) The use of embedded regular topologies reduces the average number of processing stages a packet has to traverse in the network and provides inherent load balancing.

With a smaller number of stages, the number of service instances per packet is reduced. Thus, the total number of packets processed in the network per unit of time, i.e. the network capacity, is increased. The use of regular topologies leads to further reduction of buffering requirements and a potential simplification of control procedures, such as routing and congestion control. 2) The construction of the regular topologies in a virtual mode provides a practical approach for establishing and maintaining regular structures in wide area networks, which due to distance and cabling considerations are characterized by arbitrary topologies. 3) The introduction of two level switching allows to distribute the processing/switching requirements between the electronic and optical switching capabilities of WDM systems. In Lightnet, the electronic switching capability of the nodes determines the number of channels that will be processed/switched electronically, i.e. the virtual node degree. The optical switching is made feasible by the lightpath properties: a) lightpath routes are determined at the virtual topology embedding phase, allowing the use of preset passive optical switches; b) wavelengths continuity within the lightpath, allows to significantly reduce the switch complexity.

The performance results presented in the paper quantitatively support the basic objectives of the Lightnet design. They show that for diverse WAN physical topologies the user available throughput is increased up to nearly an order of magnitude.

Acknowledgments
We wish to thank A. Rosenberg for the many helpful discussions and suggestions regarding regular topology embeddings.

References

[1] S. Heatley and D. Stokesberry, "Analysis of Transport Measurements Over a Local Area Network," *IEEE Comm. Magazine*, Vol. 27, No. 6, June 1989.

[2] E. Nussbaum, "Communication Network Needs and Technologies – A Place for Photonic Switching ?" *Jour. Sel. Areas in Comm.*, Vol. 6, No. 7, August 1988.

[3] I. Chlamtac and W. R. Franta, "Rational, Directions and Issues Surrounding High Speed Computer Networks," *Proceedings of the IEEE*, January 1990.

[4] I. Chlamtac, A. Ganz and G. Karmi, "Tree Based Lightnet Topologies," *Tech. Report # CSE-90-10, Dept. of ECE*, Univ. of Massachusetts, February 1990.

[5] I. Chlamtac, A. Ganz and G. Karmi, "Purely Optical Networks for Terabit Communications," *Infocom '89*, Ottawa, April 1989.

[6] A. Rosenberg, Personal Communication, February, 1990.

[7] M.R. Garey and D.S. Johnson "Computers and Intractability : A Guide to the Theory of NP – Completeness," *W. H. Freeman & Company* 1979.

[8] Y. Gurevich and S. Shelah, "Expected Computation Time for Hamiltonian Path Problem," *SIAM Jour. on Comp.*, Vol. 16, No. 3, June 1987.

[9] S. Even, "Graph Algorithms," *Computer Science Press Inc.*, 1979.

[10] I. Chlamtac, A. Ganz and G. Karmi, "Lightnets : Topologies for High Speed Networks," *Tech. Report # CSE-90-9, Dept. of ECE*, Univ. of Massachusetts, February 1990.

[11] A. Rosenberg, "Encoding Data Structures in Trees," *Jour. of the ACM*, Vol. 26, No. 4, October 1979.

[12] I. Chlamtac, A. Ganz and G. Karmi, "The Lightpath Architecture," *Tech. Report # CSE-90-8, Dept. of ECE*, Univ. of Massachusetts, February 1990.

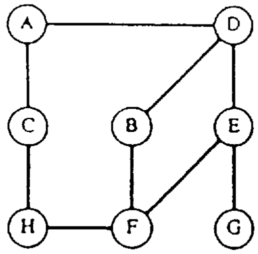


Figure 1(a): Sample General Topology Network

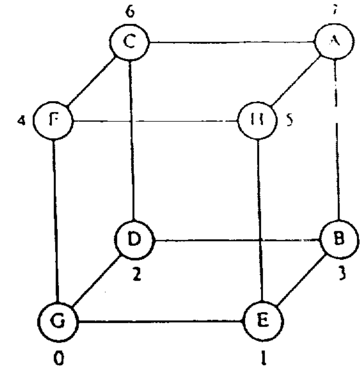


Figure 1(c): An Eight Node Hypercube

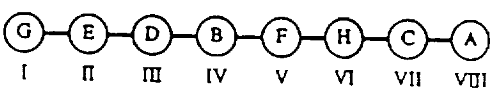


Figure 1(b): String Representation

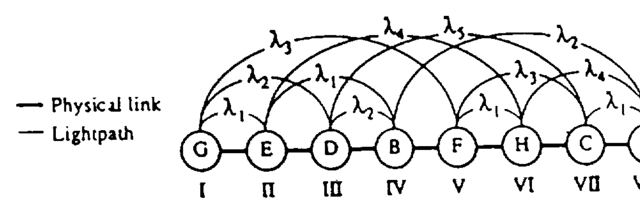


Figure 1(d): Embedding of an 8 Node Hypercube in a String

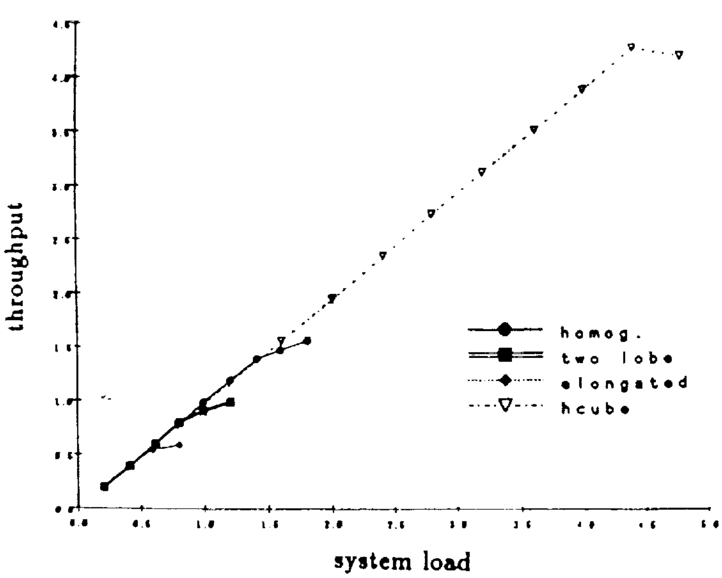


Figure 2: Throughput vs. Load for Store and Forward and Lightnet Operation

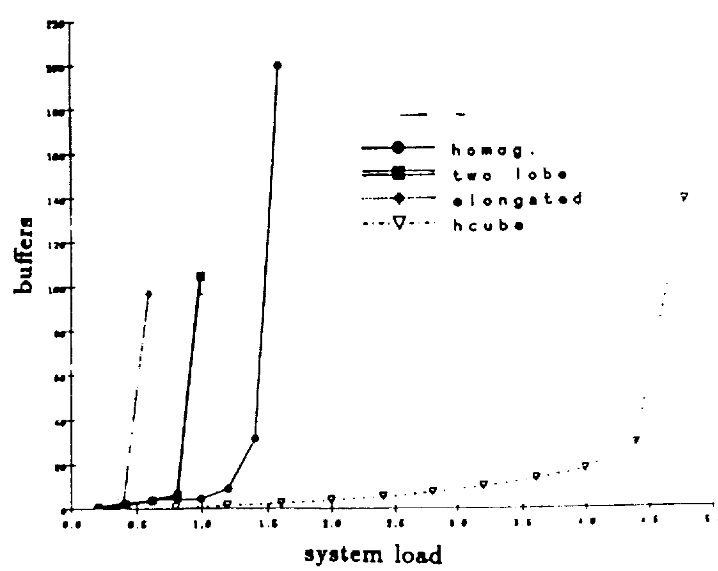


Figure 3: Buffering Requirements vs. Load for Store and Forward and Lightnet Operation

ORIGINAL PAGE IS
OF POOR QUALITY

APPENDIX C

presented at

Infocom '89
Ottawa, Canada

Purely Optical Networks for Terabit Communication

I. Chlamtac A. Ganz G. Karmi
University of Massachusetts
Amherst MA 01003

Abstract

Emerging applications require a substantially higher bandwidth than the one offered by current networks. The technology necessary for providing the high bandwidth on the optical fibers, by means of Wavelength Division Multiplexing (WDM) exists. However, none of the network architectures proposed so far can efficiently tap this bandwidth, due to the limitations imposed by the processing, buffering and switching required in these solutions. In this paper we propose a novel architectural approach that meets the high bandwidth requirements by introducing a communication architecture based on *lightpaths* – purely optical transmission paths in the network. Since lightpaths form the building block of the proposed architecture, its performance hinges on their efficient establishment and management. We show that although the problem of optimally establishing lightpaths is NP-Complete, distributed heuristics provide near optimal solutions for several of the basic problems motivated by the Lightpath architecture.

1. Introduction

Current network architectures fail to meet the emerging integrated demands of communication applications. First and foremost, a substantial increase in network bandwidth must be provided to support applications such as HDTV, super–computer communications and video–conferencing [1–9]. Co–existing with these vast bandwidth consumers, there will continue to be applications with substantially smaller requirements. Thus, in addition to the need for high bandwidth, a bandwidth dynamic range of up to seven orders of magnitude must be contended with *efficiently* [21]. Reliability and availability will also become critical issues in future high speed networks carrying services previously supported by different networks. Clearly, the degree of reliability of the new network must be at least as high as that provided in the past by the network carrying the most stringent of the integrated applications. Finally, many of the emerging applications will present demands both for predictable service and on demand data delivery, leading to the requirement for integrating packet and circuit switched policies on the same network.

Currently, Wavelength Division Multiplexing (WDM) [16–18] offers a solution to the problem of *transmitting* the required bandwidth on optical links. However, the existing switching, processing and buffering technologies lag behind the transmission capabilities, turning the nodes into the foci of congestion. Therefore the bandwidth provided by optical communication links cannot be readily translated into a user available bandwidth.

The leading approaches for wide bandwidth WANs are solutions based on packet switching, usually termed "fast packet switching" (also ATM, ATD) [22–26]. In these solutions packets are not required to wait and be error checked at intermediate nodes. However, E/O conversion of the packet header and routing oriented processing are required and in case the outgoing link is busy, the packet is either stored or discarded. Packet switching solutions are inherently characterized by efficient utilization of bandwidth at the expense of increased processing in the nodes. Therefore the node bottlenecks created by the discrepancy between transmission and processing/buffering capabilities are not removed, leading to networks with insufficient bandwidth and unbounded delays.

In this paper we propose an innovative solution to the problem of supplying wide bandwidth to the users. We employ WDM not only to attain the required bandwidth but also to simplify switching.

The use of WDM for switching purposes, is strongly motivated by considering current time division multiplexing standards for high speed WAN communication [10,11]. In these, the inherent correspondence between time slots and data channels is utilized so that no identification of the packet header is required at intermediate nodes. This leads to practical and simple switching without the need for processing. Analogously, WDM possesses the inherent capability to identify data channels without processing, through the association of these with wavelengths.

Consistently with these observations the proposed architecture is based on the use of *lightpaths*. A lightpath is an all optical path (data channel) established between any two nodes in the network, created by the allocation of the same wavelength throughout the path. A lightpath requires no E/O conversions, processing or buffering at intermediate nodes. Therefore the lightpath concept removes the electronic bottlenecks allowing efficient utilization of the bandwidth made available by WDM while reducing cost and increasing the network reliability.

Employing lightpaths as the sole medium for all network communications thus presents significant advantages. However, due to the limited number of available wavelengths, it is generally not possible to establish a lightpath between every pair of nodes in the network. In order to efficiently utilize the available wavelengths, a lightpath network, *Lightnet*, is established. The Lightnet nodes correspond to the actual network nodes while the links correspond to the established lightpaths. The Lightnet topology has as objective the minimization of the number of nodes actively involved in transmitting a packet, therefore minimizing the processing/buffering required to transmit a packet end to end. This topology can be further optimized for routing, congestion or special reliability requirements. By using the lightpaths as the common and only transmission medium for packet and circuit switched communication, the proposed architecture provides the required high bandwidth and offers integrated on-demand and connection-oriented data transmission.

Since the lightpath is the basic building block, its correct and efficient establishment is crucial to the successful implementation of the Lightpath architecture. Following an overview of the architecture we therefore proceed to study this problem in detail, analyzing its complexity and providing distributed solutions for it.

2. The Lightpath Architecture

We introduce the *lightpath* as a "direct communication path" between two (not necessarily adjacent) nodes, established by allocating the *same* wavelength throughout the route of the transmitted data. As a result, transmissions between lightpath endpoints require no electro–optic conversion and no processing at intermediate nodes. Therefore, lightpath communication can be readily implemented in an all–optical WDM network, can be managed by pure end-to-end control and carries data at "the speed of light" across the network.

To understand why the lightpath concept and the architecture built around it are a natural development in communication and to clarify the principles of a lightpath based communication architecture we consider the following analogy :

With the increase in the speed of trains and growing congestion at stations, express train transportation systems were developed. When taking an express train, the passenger does not have to wait at intermediate (local) stations and his travel time becomes determined only by the speed of the express train. Therefore, the express system solved the congestion problem, and lead to lower passenger delays and smaller waiting room at intermediate stations.

The increased speed of optical communication contrasted with the speed of electronics at the switching nodes creates an apparent technological analogy: In our proposed solution, the wavelengths are the rails, the lightpath is an express connection established between two stations and circuits/packets are the trains. Furthermore, just as in railways, the number of wavelengths (rails) is not sufficient to establish a lightpath (express connection) between any two stations in the network. The limitation on the number of lightpaths that can be established thus makes the efficient establishment of lightpaths a crucial issue.

Due to the need for end–to–end lightpath establishment, lightpaths cannot meet the needs for on demand communication through on demand lightpath establishment. The lightpaths are thus further equated with express train connections, as clearly, they are more dynamic than creation of wavelengths

(laying rails), but not as dynamic as changes in traffic (passenger) distribution whether in a network or a railway system. Therefore, to create an efficient communication network we carry our analogy one step further by establishing a fast communication structure, the *Lightnet*, based on lightpaths similarly to the design of an express train schedule in a railway system. The design of the Lightnet will take into consideration the number of wavelengths, the maximum number of hops possible without E/O conversions and the underlying network topology. On the basis of the common Lightnet, the proposed architecture constructs an integrated packet and circuit switching solution. For packet switching, packets will be routed over "adjacent lightpaths", using available free bandwidth, instead of being routed between physically adjacent nodes, as in conventional packet switched networks. In the proposed circuit switching solution, bandwidth over several lightpaths will be allocated to a circuit for the circuit's duration.

The presented architecture carries several fundamental benefits :

- The Lightnet reduces the number of active nodes a packet travels from source to destination, thus alleviating the performance and reliability bottlenecks, created by E/O conversions, processing and storage.

- The lightpaths present a novel approach to solving congestion problems, fault conditions and offer particularly attractive solutions to networks having asymmetric traffic patterns. The capability to account for these issues is a result of the quasi-dynamic nature of the Lightnet.

- The switching nodes' hardware requirements are simplified, enabling the use of *relational* devices in which the relation of the mapping between the inputs and the outputs is *independent* of the data [27]. A possible implementation of a switching node for the Lightpath architecture is depicted in figure 1. We refer to [13–15,27] for descriptions of suitable photonic switches.

3. The Lightpath Establishment Problem

Since the performance of the Lightpath architecture hinges on the efficient establishment of lightpaths, we turn now to study this problem in detail. We note that the requirement for wavelength continuity intuitively leads to a bandwidth loss when compared to systems where the continuity constraint is not imposed. This loss can be perceived either as an increase in the number of wavelengths required to successfully establish a given set of lightpaths, or as an increase in the blocking probability if the number of wavelengths is limited. In providing solutions for lightpath establishment, our objective will therefore be to find algorithms that minimize this loss.

In deriving a lightpath establishment algorithm, we first analyze the complexity of an optimal assignment of lightpaths, introducing the following model. We represent the network by a triplet $G(V, E, W)$ in which V represents the set of nodes, E represents the set of directional fiber links between nodes in V, and W is the set of wavelengths on each link, $|W| = \omega$. We shall assume that ω is equal for all links.

Definition : A *lightpath request* is defined by the links constituting the lightpath that has to be established between a source and a destination node. For transmissions to proceed on a lightpath, the lightpath must be *established* by finding an unallocated, identical, and properly setting up the photonic space switches at the intermediate nodes.

The problem we propose to study is the *correct* and *efficient* establishment of lightpaths. The correctness aspect of lightpath establishment must solve the problems of collisions : the simultaneous allocation of the same wavelength to more than one lightpath on any given link. In terms of efficiency, our goal is to maximize the utilization of wavelengths. Thus we shall seek solutions that minimize lightpath blocking probability, as a function of lightpath requests. We propose to achieve this goal by allocating resources in such a way that, given the allocation of wavelength to existing lightpaths, a maximal number of new lightpaths can be allocated. Figures $2a$ and $2b$ exemplify the lightpath allocation problem. The figures depict lightpaths establishment in a network where WDM is employed, with two available wavelengths ($\omega = 2$). In figure $2a$ the allocation is done in such a way

that any (single) future lightpath demand can be allocated. In the allocation depicted by figure $2b$, if a lightpath demand $v_1 \rightarrow v_3$ comes up before an existing lightpath is terminated, it will be blocked.

We next consider a lower bound on the complexity of the above "Dynamic Lightpath Establishment" (DLE) problem by showing that a simpler, closely related problem is NP–Complete :

Definition : Static Lightpath Establishment (SLE) problem – given a network $G(V, E, W)$, $\omega \geq 3$, and a predefined set of lightpaths L, is it possible to establish all lightpaths in the set ?

We proceed to prove the NP-completeness of SLE by showing that the problem is equivalent to the n–graph–colorability problem [19,20]. That is, finding the minimal number of resources that would accommodate the demands would amount to finding the chromatic number of some (general) graph, where the number of colors, n, corresponds to the number of wavelengths, ω.

Theorem : SLE is NP-Complete.

Proof : First we show that solving the n–graph–colorability problem would also solve SLE. Define an undirected graph $G_L(V_L, E_L)$ with a node $v \in V_L$ for every lightpath in L. Two vertices $v_1, v_2 \in V_L$ have an interconnecting edge $e \in E_L$ if the respective lightpaths have at least one link in common. A coloring of V_L with n or less colors, so that no two adjacent vertices have the same color, would yield a resource allocation in W where no two lightpaths having a link in common require the same resource. Thus, finding a feasible coloring would also yield a feasible resource allocation, answering SLE.

To complete the proof we show that solving SLE would also solve the n–graph–colorability problem, thus showing that finding a polynomial solution to SLE is unlikely. To show this, we describe a polynomial time algorithm that translates any graph into a network and an appropriate set of lightpath demands. Given a graph $G_L(V_L, E_L)$ do :

(1)create a node v_i^0 for every node $i \in V_L$.
(2)for every edge $e = i \rightarrow j \in E$:
create 4 new nodes x, y, v_i^k, v_j^l
and directed edges
$v_i^{k-1} \rightarrow x, \quad v_j^{l-1} \rightarrow x, \quad x \rightarrow v_i^k, \quad x \rightarrow v_j^l$
Attach the mark i to edges going
from/to v_i's, and $x \rightarrow y$.
Repeat similarly for the mark j.

The construction is exemplified for a 4 node graph in figures $3a$, $3b$. Figure $3a$ contains a graph for which the n–colorability problem is to be solved. Figure $3b$ illustrates its translation to a network, the numbers on the links being the marks. The lightpath demand set L is defined by the $|V_L|$ lightpaths where lightpath i requires use of all links having i as a mark. We note that the complexity of the algorithm is $O(|E_L|)$.

Lemma : A solution to the SLE with a resource set of size n implies that the chromatic number of G_L is less or equal to n.

Proof : The lemma follows immediately from the construction. If the lightpaths can be established then there exists a function assigning a resource to each lightpath so that no lightpaths sharing a link are assigned the same resource. Since two lightpaths share a link if and only if the respective nodes in V_L are adjacent, this implies the existence of a function assigning a color to each node in V_L, so that no two adjacent nodes are assigned the same color.

Thus, even if all lightpath demands were predetermined, we would have to search for a heuristic solution for all but trivial demand sets. When dealing with an environment were lightpaths are requested and terminated dynamically, an efficient lightpath establishment becomes even more difficult as future demands cannot be predicted. The next section presents a number of polynomial time solutions for dynamic lightpath establishment. Surprisingly, these demonstrate that relatively simple heuristics can yield very good results.

4. Lightpath Establishment Heuristics

In searching for heuristic solutions, our purpose is twofold : investigate the performance of each heuristic and, obtain heuristics that produce correct, efficient allocations distributively for dynamic lightpath demands. The case we consider is that of non–alternating, source routing. As computing an optimal allocation is intractable for any traffic demand set of interest, the problem of evaluation becomes also difficult. The translation to the

colorability problem could have been theoretically used for comparison to lower bounds on the chromatic number. Unfortunately, no good bounds are known. Furthermore, the best bound, $\gamma \geq n/\alpha$, where n is the number of nodes, α the stability number and γ the chromatic number, is both expensive to compute and can be shown to be arbitrarily bad [19,20].

We shall therefore study the performance of the distributed heuristics from two perspectives. The first is concerned with the "performance penalty" of the lightpath approach and the continuity constraint thereof in terms of blocking probabilities. To obtain this, we perform a comparison with "conventional circuit switching" where lightpaths are established for as long as there is *any* resource available on each link, i.e. not necessarily the same one. Such would be the case if "ideal" wavelength convertors with zero delays were available. The second issue is to compare the results obtained in a distributed way with those that can be obtained if all information is available, in a centralized way. In what follows, we first present a Centralized Lightpath Allocation (CLA) heuristic, compare it to to the "ideal" case and then proceed with the presentation of distributed heuristics and their performance, relative to CLA.

4.1 Centralized Lightpath Allocation

In order to minimize lightpath blocking probability we first consider an approach based on the principle of achieving maximal wavelength reuse throughout the network. The intuition supporting this approach is twofold :

1. As long as there is at least one wavelength λ_i which is not allocated on any link in the network, we are guaranteed that any new lightpath demand can be met with no blocking.

2. Assume that a given wavelength λ_i has already been allocated in a subset $E_i \in E$. The larger this subset, the smaller is the proportion of new lightpaths which can be established allocating λ_i. Thus, for any new lightpath demand that can be established, using one of $\lambda_1, \lambda_2, \ldots$, we should perform this allocation by assigning it the wavelength λ_i (in the group) with largest E_i set.

We next present a formal solution based on this approach.

The data structures used in CLA are :

l : lightpath demand; array containing the links defining the lightpath.

wave : array determining wavelength utilization; *wave*$[i]$ is the number of links in which wavelength i is allocated.

alloc : 0 / 1 matrix; *alloc*$[i, j] = 1$ if wavelength j is allocated in link i (hence *wave* is the sum of the columns of *alloc*).

The heuristic scans *alloc* finding which wavelengths are feasible, choosing among them, the one for which *wave*$[i]$ is maximal :

CLA algorithm

```
establish(l)
(* establish a lightpath l *)
  max = -1
  for i = 1 to n begin (* for all wavelengths *)
    feasible = 1
    for all links k on l's path
      feasible = feasible and alloc[l[i],k]
    if feasible and if max < wave[i] then
      max = wave[i], w = i
  end
  if max > 0 then begin
    for i = 1 to len do alloc[l[i],w] = 1
    wave[w] = wave[w] + len
    "establish" the lightpath
  end
  else - the lightpath is blocked

terminate(l,w)
(* terminate a lightpath l established
  by using wavelength w *)
  for i = 1 to l do alloc[l[i],w] = 0
  wave[r] = wave[r] - len
```

The above heuristic was evaluated using a simulation with the following parameters : Lightpath duration times were taken as constant (200 time units) and lightpath inter–arrival time as exponential. All results were measured with a confidence level of 99%. Traffic was assumed to be uniformly distributed; routing was non–alternate, shortest path, choosing a path at random when several were possible. 10 wavelengths were assumed to be available on each link. In figures 5–10 blocking probabilities

are given as a function of lightpath arrival rate measured in lightpaths per time unit. Figure 4*a* depicts a sample general topology network. Figure 5 depicts the blocking probabilities for CLA and conventional circuit switching averaged over all lightpaths and for longest lightpaths only. In terms of average blocking probability we see that the results are very close and in fact, for certain loads, CLA actually displays slightly lower blocking probabilities. This is due to the fact that long lightpaths are rejected by CLA with a higher probability than in conventional circuit switching (see figure 5). The higher rejection probability occurs due to the requirement to find an identical free wavelength throughout the path. Hence since with CLA more short lightpaths will be established, the average blocking probability is decreased.

4.2 PACK – A Distributed Heuristic

As pointed out in section 2, for a realistic implementation of the lightpath approach, distributed heuristics are needed. Lacking the global information used in CLA to maximize resource re-use across the network, two viable approaches for distribution can be taken. The first approach is based on exchanging information between neighbors, eventually creating a global picture, or an assessment thereof, in each node. This approach is useful when the lifetime of the information is long with respect to the information propagation time. However, when the structures described have connection times that may be short, nodes will be making decisions based on outdated information most of the time. We further point out that this approach also incurs an additional complexity cost in bandwidth dedicated to control. The second alternative is to emulate global knowledge by implementing a global policy. This can be done by requiring that nodes that decide which resource will be allocated to a lightpath do so, by using the same rules. This is precisely what the following "PACK to beginning" heuristic does. Let $\lambda_1, \lambda_2, \ldots$ be any arbitrary numbering of the wavelengths known to all nodes. PACK will allocate the smallest numbered wavelength feasible. Thus on a new lightpath requirement PACK emulates CLA in the attempt to maximize re-use of wavelengths allocating them in the same order in all nodes. However, CLA adapts this order according to the current wavelength allocation while PACK uses a fixed, preset ordering. Hence, CLA has a superior ability to adapt to changes in lightpath demands. Performance discrepancies between these two heuristics may therefore be expected due to this difference.

In the PACK distributed solution four types of messages are exchanged between the nodes. The message length is, in the worst case, $O(\omega)$. These messages are :

REQUEST (src,dest,wave,id) : lightpath establishment request. *wave* is a bitvector containing a "0" in the i'th location if wavelength i can be allocated for the lightpath. *id* is a unique lightpath identifier, obtained locally by concatenating the originating node id to some counter.

ACCEPT (src,dest,i,id) : lightpath establishment notice. i is the wavelength number allocated to the lightpath.

REJECT (src,dest,id) : rejection notice issued when a lightpath request is blocked.

HANGUP (src,dest,id) : lightpath termination message, initiated by node originating the lightpath request.

Each node maintains the following data structures :

lightpath(id) An array containing a record for each lightpath passing through the node. The record contains the wavelengths allocated (or reserved) for the lightpath, its incoming edge and its outgoing edge.

switch$[1..n, 1..d, 1..d\ ]$ where $\omega = n$ and d is the degree of the node. **switch**[] defines the wavelength allocation and the appropriate switching function in the node (e.g. $switch[3,2,4] = 1$ indicates that link 2 is to be switched to link 4 for wavelength λ_3).

Following is the algorithm executed by each node upon receipt of the corresponding messages :

PACK algorithm

request(src,dest,wave,id)
    if dest = node then begin
        i = select slot(wave)
        accept(src,dest,i,id)
    end
    else begin
        n = next_ node(dest) (next node in route *)*

```
        for every wavelength i
        if i used in incoming/outgoing link then
            wave[i] = 1
        if wave[i]=1 ∀i reject(src,dest,id)
        else begin
            update data structures for lightpath id
            if wave[i] set to '1' in this node,
                set to '1' the relevant entry in switch
            send(n,REQUEST,node,dest,wave,id)
        end
    end
end

accept(src,dest,i,id)
    set to '0' all entries previously set to '1'
        in switch for id, except i
    let n be the incoming node of id
    (from lightpath(id))
    if src ≠ node send(n,ACCEPT,src,node,i,id)
end

reject(src,dest,id)
    set to '0' all entries previously set to '1'
        in switch for id
    let n be the incoming node of id
    (from lightpath(id))
    if src ≠ node send(n,REJECT,src,node,id)
end

terminate(src,dest,id)
    if node ≠ dest begin
        free entry in switch corresponding to id
        let n be the outgoing node of id
    (from lightpath(id))
        send(n,TERMINATE,node,dest,id)
    end
end
```

where *select_ slot* return the lowest numbered feasible wavelength.

Figure 6 contains a comparison in terms of blocking probabilities for PACK and CLA for the sample network. As can be seen, the results are practically identical, both for the average length lightpath blocking probability and the longest lightpath blocking probability. Hence, by transmitting information only along the path of the lightpath, we have obtained, contrary to intuition, a distributed heuristic paying a negligible price in terms of performance.

4.3 Special Topologies

To better understand the relative performance of the CLA and PACK heuristics we considered additional network topologies. Figure 4*b* depicts a network where the topology and traffic pattern create a single link bottleneck. Blocking probabilities for this network, shown in Figures 7 and 8, analogous to figures 5 and 6, demonstrate that earlier observations remain valid. This is explained by the fact that in the general topology network chosen in the previous case, bottlenecks are also bound to occur (although at different points in the network at different times). It is thus of interest to consider the other extreme network case for relative performance of the two heuristics. Namely we want to consider a completely symmetrical network, a torus, in which no single bottleneck will occur, as depicted in figure 4*c*. The differences between the performance of the two heuristics are in this case more pronounced, as can be seen from figures 9 and 10. A maximal discrepancy, 15%, occurs between PACK and CLA for the longest lightpaths. The quantitative difference in the behavior of the two algorithms as a functions of the network topologies can be explained from the following observation. In the presence of bottlenecks, long lightpaths cannot be established, not even under conventional circuit switching. Hence, in these networks the importance of optimal wavelength reuse is less pronounced than in symmetric networks. Therefore, in symmetric networks, the prime factor determining the establishment probability is the wavelength allocation method.

5. Conclusions

In this paper we presented a novel network architecture motivated by recent developments in optical communications and targeted towards emerging wide bandwidth applications. The architecture makes use of developing transmission and switching capabilities in the photonic domain to overcome the inherent limitations of electronics by introducing the lightpath concept. It was shown how, based on lightpaths, an architecture meeting the requirements of emerging wide bandwidth applications, can be designed. Since the performance of this architecture is linked, first and foremost, to the efficient establishment of lightpaths, a detailed investigation of the lightpath establishment problem was conducted. The complexity of this problem was

studied and proven to be NP–Complete. However, it was shown, that using polynomial time heuristics near–optimal results can be obtained. Both centralized and distributed establishment heuristics for the dynamic lightpath establishment problem were presented.

References

1. J. O. Limb, "Performance of Local Area Network at High Speed," *IEEE Communication Magazine*, Vol. 22, No. 8, August 1984.

2. S. Personick, "Review and Fundamentals of Fiber Optics Systems", *IEEE Journal on Sel. Areas in Comm.*, Vol. SAC-1, No. 3, April 1983.

3. P. Henry, "Introduction to Lightwave Transmission", *IEEE Communications Magazine*, May 1985.

4. D. Keck, "Fundamentals of Optical Waveguide Fibers", *IEEE Communications Magazine*, May 1985.

5. E. Basch and T. Brown, "Introduction to Coherent Optical Fiber Transmission", *IEEE Communications Magazine*, May 1985.

6. J. S. Turner, "New Directions in Communications" *IEEE Communications Magazine*, October 1986.

7. M. R. Finley, "Optical Fibers in Local Area Networks", *IEEE Communications Magazine*, August 1984.

8. P. Kaiser, J. Midwinter and S. Shimada, "Status and Future Trends in Terrestrial Optical Fiber Systems in North America, Europe and Japan," *IEEE Communications Magazine*, Vol. 25, No. 10, October 1987.

9. R.J. Boehm, Y.C. Ching, C.G. Griffith and F.A. Saal, "Standardized Fiber Optic Transmission Systems", *IEEE Journal on Sel. Areas in Comm.*, Vol. SAC-4, No. 9, December 1986.

10. T1X1.4 contribution, "Rates and Formats for Optical Network Interfaces",T1X1.4/85–017.

11. T1X1.4 contribution, "Synchronous Optical Network (SONET)", T1X1.4/85–005.

12. M. Schwartz, "Telecommunication Networks – Protocols, Modeling and Analysis", Addison Wesley Publishing Company, 1987.

13. T. Sawano, S. Suzuki, M. Kondo, K. Nagashima, "Photonic Switching System for Broadband Services", *Globecom 1987*

14. A. M. Hill, "One Sided Rearrangeable Optical Switching Networks", *Journal Of Lightwave Technology* Vol. LT-4, No. 7 July 1986.

15. A. M. Hill, "Switching and Distribution Networks for Wideband Optical Signal", *IEEE Communication Society International Switching Symposium* March 1987.

16. G. R. Hill, "A Wavelength Routing Approach to Optical Communications Networks", *Globecom 1988* Tokyo, Japan, 1988.

17. H. Ishio, "Review and Status of WDM Technology and its Applications", *Journal on Lightwave Technology,* Vol LT-2, No. 4, August 1984.

18. M. C. Brain, P. Cochrane, "Wavelength Routed Optical Networks Using Coherent Transmission", *ICC88*, Philadelphia, June 1988.

19. Garey, D.S. Johnson "Computers and Intractability : A Guide to the Theory of NP – Completeness", *W. H. Freeman & Company* 1979.

20. N. Christofides, "Graph Theory – An Algorithmic Approach", *Academic Press* 1975.

21. E. Nussbaum, "Communication Network Needs and Technologies - A Place for Photonic Switching?", *IEEE Journal on Sel. Areas in Comm. joint with Lightwave Technology*, Vol. 6, No. 7, August 1988.

22. Z. Haas, D.R. Cheriton, "A Case For Packet Switching in High – Performance Wide Area Networks", *SIGCOMM 87'* Vermont, Aug 1987.

23. J. Turner, "Design of an Integrated Services Packet Network", *IEEE Journal on Sel. Areas in Comm.*, Vol. 4, No. 8, Nov. 1986.

24. J. Coudreuse, M. Servel, "Prelude : An Asynchronous Time Division Switched Network", Proc. 1987 IEEE Int. Conf. on Comm., 1987.

25. P. Newman, "A Fast Packet Switch for the Integrated Services Backbone Network", *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 6, Dec. 1988.

26. H. Suzuki, T. Takeuchi, F. Akashi, T. Yamaguchi, "Very High-Speed and High-Capacity Packet Switching for Broadband ISDN", *IEEE Journal on Sel. Areas in Comm.*, Vol. 6, No. 6, Dec. 1988.

27. H.S. Hinton, "Photonic Switching Systems", *ECOC 1988*, Sept. 1988.

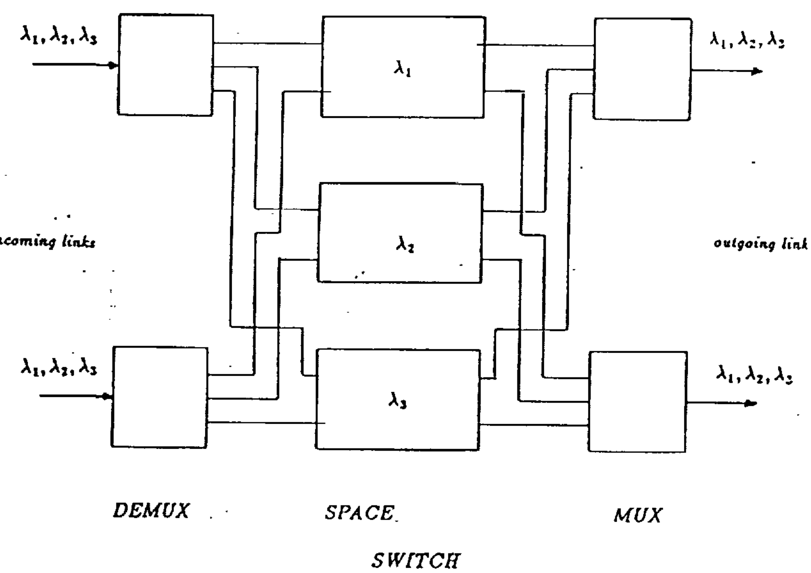


Figure 1: WDM implementation of a switching node.

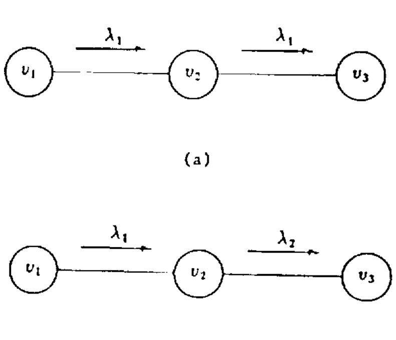


Figure 2: Examples of lightpath allocation.
L_1: v_1 to v_2
L_2: v_2 to v_3

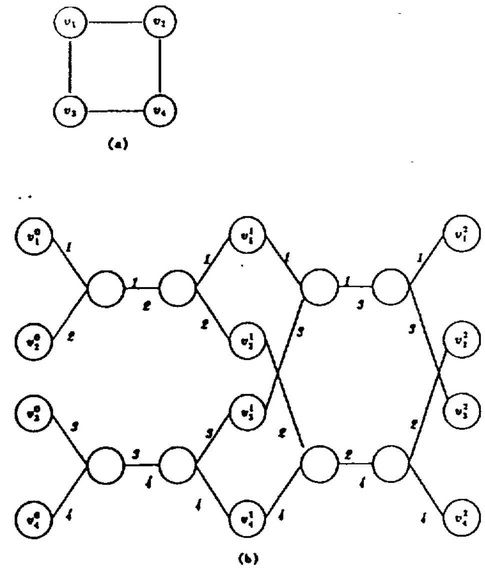
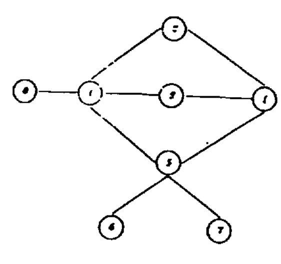
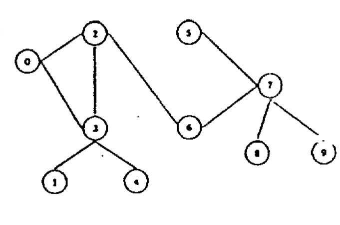


Figure 3: (a) n-colorability graph.
(b) Translation to SLE.





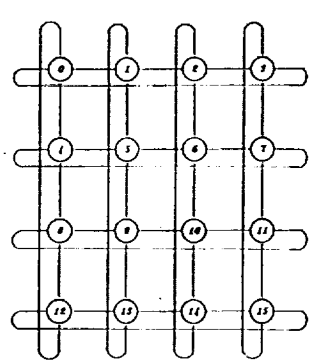


Figure 4: Sample Topologies
(a) general
(b) bottleneck
(c) torus

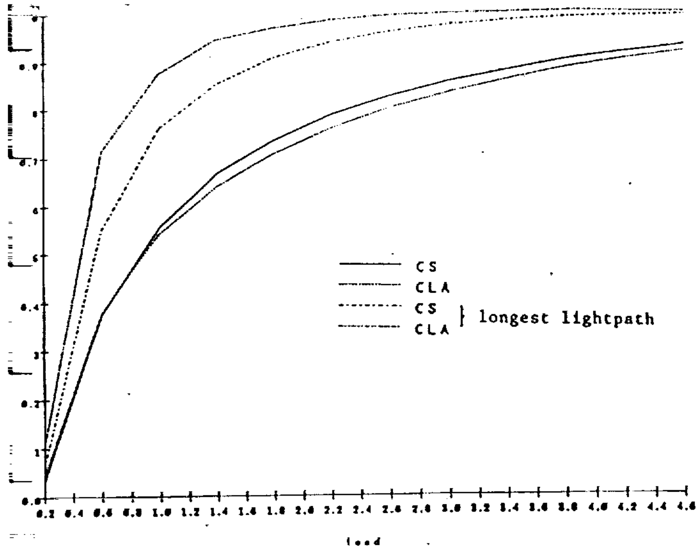


Figure 5: Blocking probabilities for CLA and CS algorithms for the network in Figure 4a.

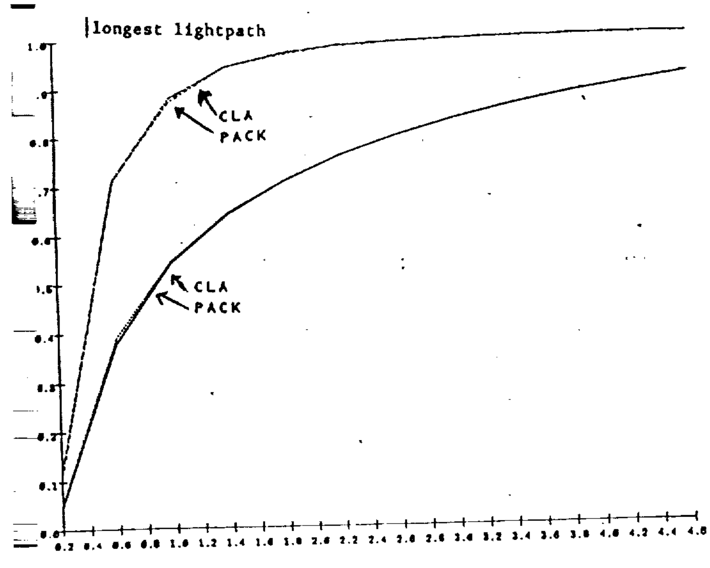


Figure 6: Blocking probabilities for CLA and PACK algorithms for the network in Figure 4a.

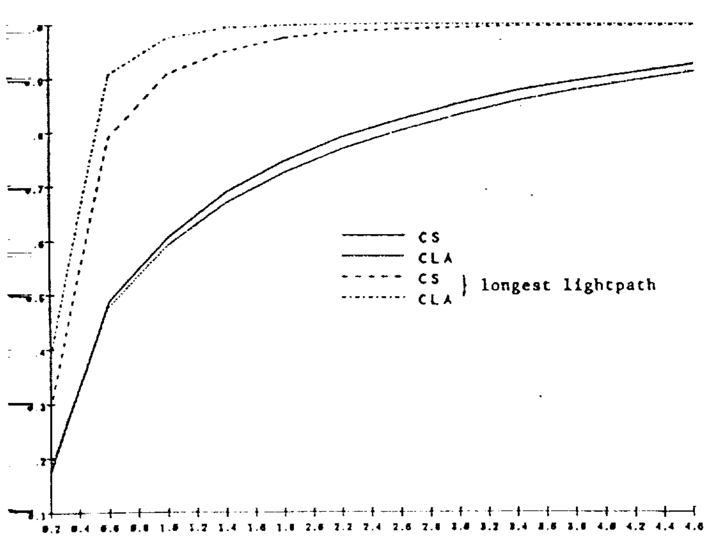


Figure 7: Blocking probabilities for CLA and CS algorithms for the network in Figure 4b.

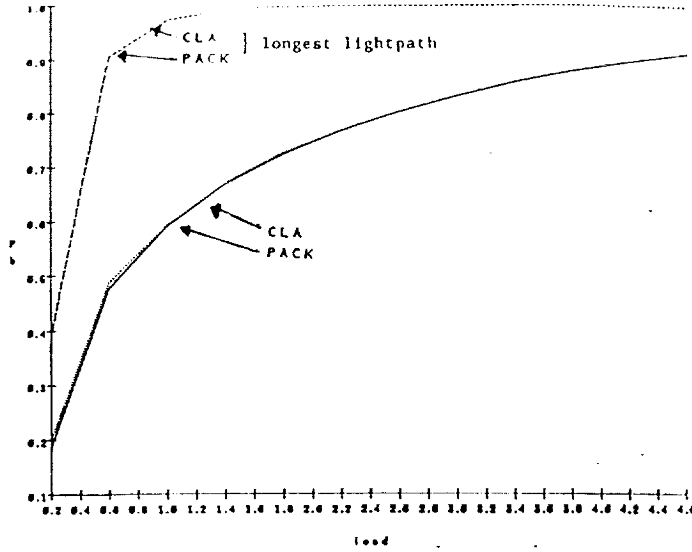


Figure 8: Blocking probabilities for CLA and PACK algorithms for the network in Figure 4b.

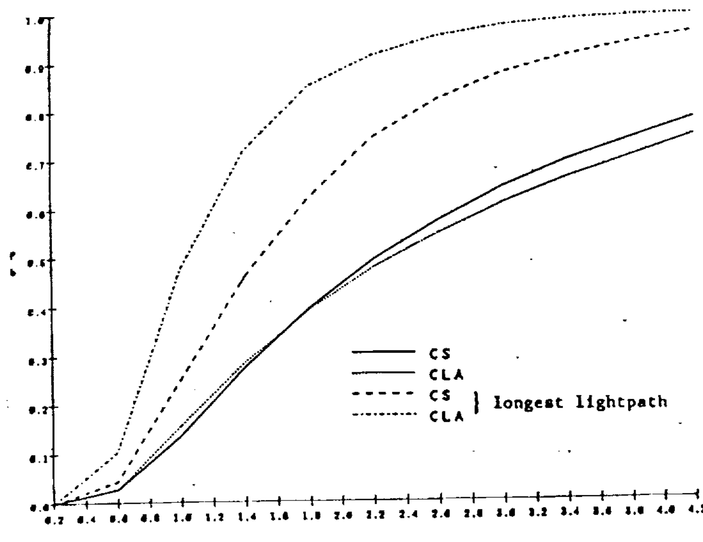


Figure 9: Blocking probabilities for CLA and CS algorithms for the network in Figure 4c.

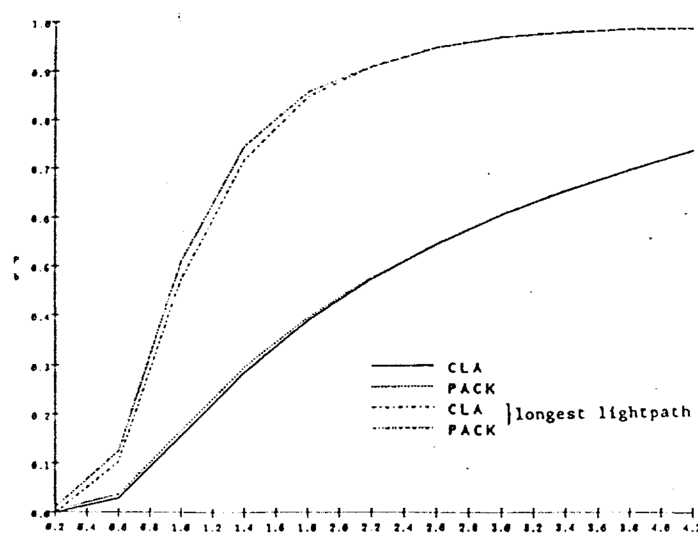


Figure 10: Blocking probabilities for CLA and PACK algorithms for the network in Figure 4c.